

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-110,115
(43)Date of publication of application : 22.04.1994

(51)Int. Cl.

G06F 9/46

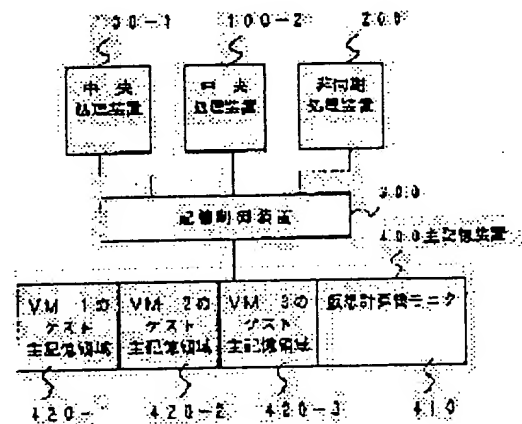
(21)Application number : 04-257013
(22)Date of filing : 25.09.1992

(71)Applicant : HITACHI LTD
(72)Inventor : TANAKA TOSHIHARU
ARAI TOSHIAKI
SUGAMA NOBUYOSHI
YAMAMOTO TADASHI
IKEGAYA NAOKO

(54) DYNAMIC ALLOCATING METHOD FOR COMPUTER RESOURCES IN VIRTUAL COMPUTER SYSTEM

(57)Abstract:

PURPOSE: To provide a virtual computer system which is dynamically changing the allocation of a main storage area to an OS (operating system).
CONSTITUTION: With the occurrence of a specific event as an opportunity, a high-order address area within the guest main storage area 420-2 of a virtual computer VM-2 is made to an OFF line (unconnection state). Then, a virtual computer monitor 410 moves the guest main storage area 420-2 of the VM-2 and allows the high-order address area 420-1 of a virtual computer VM-1 to be an empty state. Subsequently, the high-order address area which is adjacent to the guest main storage area 420-1 of the VM-1 is made in ON-line (connection state) and is connected with the VM-1. Thus, the allocation of a main storage area to an OS is dynamically changed.



LEGAL STATUS

[Date of request for examination]
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

Copyright (C): 1998, 2000 Japanese Patent Office

特開平6-110715

(43)公開日 平成6年(1994)4月22日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 9/46	3 5 0	8120-5B		

審査請求 未請求 請求項の数9(全14頁)

(21)出願番号 特願平4-257013

(22)出願日 平成4年(1992)9月25日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 田中 俊治

神奈川県川崎市麻生区王禅寺1099 株式会社日立製作所システム開発研究所内

(72)発明者 新井 利明

神奈川県川崎市麻生区王禅寺1099 株式会社日立製作所システム開発研究所内

(72)発明者 須釜 延芳

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア開発本部内

(74)代理人 弁理士 富田 和子

最終頁に続く

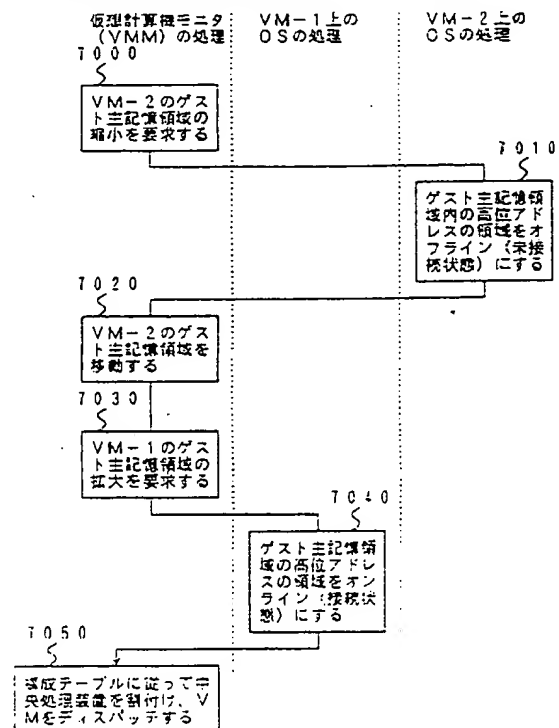
(54)【発明の名称】 仮想計算機システムにおける計算機資源の動的割付け方法

(57)【要約】

【目的】 OSへの主記憶領域の割付けを動的に変更可能な仮想計算機システムを提供する。

【構成】 特定のイベントの発生を契機として、仮想計算機VM-2のゲスト主記憶領域内の高位アドレスの領域をオフライン(未接続状態)にする(ステップ7010)。次に、仮想計算機モニタ410は、VM-2のゲスト主記憶領域を移動(ステップ7020)して、仮想計算機VM-1のゲスト主記憶領域に隣接する高位アドレスの領域を空き状態とする。その後、VM-1のゲスト主記憶領域に隣接する高位アドレスの領域をオンライン(接続状態)にして、VM-1に接続する(ステップ7040)。このようにして、OSへの主記憶領域の割付けを動的に変更する。

計算機資源の割り当て変更方法のフローチャート(図7)



【特許請求の範囲】

【請求項 1】中央処理装置と、主記憶装置と、前記中央処理装置からの指示に応じて当該中央処理装置と非同期に処理を行う非同期処理装置とを含む計算機上で複数のオペレーティング・システム（OS）が仮想計算機モニタの制御下で走行する仮想計算機システムにおいて、前記複数の OS のうち、第 1 の OS に対して、前記主記憶装置の第 1 の固定領域を割り付けるとともに、第 2 の OS に対して前記第 1 の固定領域のアドレス高位側に隣接する第 2 の固定領域を割り付けるステップと、特定のイベントが発生した時に、前記第 2 の OS に割り当てられた第 2 の固定領域のサイズを縮小するステップと、該サイズを縮小した第 2 の固定領域を前記主記憶装置上でアドレス高位側へ移動することにより、前記第 1 の固定領域のアドレス高位側に空き領域を生成するステップと、前記第 1 の OS の主記憶領域のサイズを拡大するステップとを有する、仮想計算機システムにおける計算機資源の割り付け方法。

【請求項 2】請求項 1 において、前記特定のイベントの発生前と発生後における各々の OS の主記憶領域のサイズを定義するステップと、前記特定のイベントが発生したとき、前記定義に従って前記複数の OS の主記憶領域のサイズを変更するステップとを有する、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 3】請求項 1 において、前記中央処理装置を複数個有し、特定のイベントの発生前と発生後において各々の OS に割り付ける中央処理装置を定義するステップと、前記特定のイベントが発生したとき、前記定義に従って前記複数の OS に中央処理装置を割り付けるステップとを有する、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 4】請求項 1、2、または 3 において、前記特定のイベントは、特定の時間になったというイベントであることを特徴とする、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 5】請求項 1、2、または 3 において、前記特定のイベントは、前記計算機と異なる第 2 の計算機に障害が発生し、該第 2 の計算機の処理を前記第 1 の OS が引き継いだというイベントであることを特徴とする、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 6】請求項 1 の前記空き領域を生成するステップは、前記第 2 の OS の主記憶領域を前記仮想計算機モニタが生成した仮想空間とし、該仮想空間を前記縮小された第 2 の固定領域にマッピングするステップと、前記仮想空間のマッピング先を前記主記憶装置の移動先の領域に変更するステップと、該移動先の領域を前記第 2 の OS の固定領域として割り付けるステップとを有するこ

とを特徴とする、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 7】請求項 6 において、前記第 2 の OS の主記憶領域の移動中は、前記第 2 の OS が発行した前記非同期処理装置の使用開始命令を、前記仮想計算機モニタがシミュレーションするステップと、該シミュレーション時に前記非同期処理装置が前記第 2 の OS の主記憶領域のうちどの領域をアクセスするかを解析するステップと、該アクセス対象となる領域は、前記非同期処理装置によるアクセス終了後、移動するステップとを有することを特徴とする、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 8】中央処理装置と、主記憶装置と、前記中央処理装置からの指示に応じて当該中央処理装置と非同期に処理を行う非同期処理装置とを含む計算機上で複数のオペレーティング・システム（OS）が仮想計算機モニタの制御下で走行する仮想計算機システムにおいて、前記いずれかの OS の主記憶領域として割り付けられた前記主記憶装置の連続アドレスの固定領域を移動するとき、当該 OS の主記憶領域と前記固定領域との対応関係を前記仮想計算機モニタが生成した仮想空間に定義するステップと、前記仮想空間のマッピング先を前記主記憶装置上の移動先の連続アドレスの固定領域に変更するステップと、該移動先の固定領域を前記 OS の主記憶領域として割り付けるステップとを有する、仮想計算機システムにおける計算機資源の動的割り付け方法。

【請求項 9】請求項 8 において、前記定義するステップの後に、前記仮想空間のマッピング先の固定領域のサイズを前記 OS の主記憶領域のサイズよりも小さくするステップを有することを特徴とする、仮想計算機システムにおける計算機資源の動的割り付け方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、計算機システムに関し、特に、1 台の計算機上で複数の OS が走行する計算機システムにおいて、OS への計算機資源の割り付けを動的に変更するのに好適な計算機システムに関する。

【0002】

【従来の技術】1 台の計算機上で複数のオペレーティング・システム（OS：Operating System）が走行可能な計算機システムとして、仮想計算機システム（VMS：Virtual Machine System）がある。仮想計算機システムは、1 台の計算機上に論理的な計算機である仮想計算機（VM：Virtual Machine）を複数台生成し、各仮想計算機（VM）において対応する 1 つの OS の走行を可能とする。この仮想計算機システムは、使用目的の異なる複数の OS を 1 台の計算機上で運用したり、旧バージョンの OS を運用しながら新バージョンの OS のシステムシミュレーションを行ったり、1 台の計算機上で OS の

テストを並列に実行するために利用されている。仮想計算機システムにおいて、その制御プログラムである仮想計算機モニタ (VMM: Virtual Machine Monitor) は、各仮想計算機 (VM) に計算機資源を分配し、仮想計算機 (VM) のスケジューリングとディスパッチを行なう。また、仮想計算機モニタ (VMM) は、仮想計算機上の OS が発行した命令のうち計算機では直接実行できない命令、すなわち、ソフトウェアの介入を必要とする命令のシミュレーション処理等を行なう。

【0003】仮想計算機には、仮想計算機の主記憶領域 (ゲスト主記憶領域) の与え方によってページ可能 VM とページ固定 VM がある。ページ可能 VM は、特開昭57-212680号公報記載のように、仮想計算機モニタ (VMM) が生成した仮想空間 (ホスト仮想空間) 上に、ゲスト主記憶領域を常駐させる。従って、ページ可能 VM のゲスト主記憶領域は、仮想計算機モニタ (VMM) によるページングの対象となる。このページングにより、ページ可能 VM のゲスト主記憶領域のサイズは、ホスト仮想空間のマッピング先である仮想計算機モニタ (VMM) のページング領域のサイズよりも大きくできる。

【0004】他方、ページ固定 VM は、特開昭60-24735号公報記載のように、主記憶装置上の連続領域にゲスト主記憶領域を常駐させる。従って、ページ固定 VM のゲスト主記憶領域は、仮想計算機モニタ (VMM) によるページングの対象とならない。このため、ページ固定 VM は、ページ可能 VM に比べてアドレス変換に要するオーバーヘッドが小さい。しかし、ページ固定 VM のゲスト主記憶領域のサイズは、対応付けられた主記憶装置の連続領域の大きさしか設定できない。

【0005】また、特開平3-211629号公報は、中央処理装置とは非同期に処理を行なうチャネル等の非同期処理装置によるゲスト主記憶領域へのアクセスを一時停止して、主記憶装置上でページ固定 VM のゲスト主記憶領域を移動可能とする仮想計算機システムを開示している。この仮想計算機システムによれば、主記憶装置の記憶領域の断片化を、ゲスト主記憶領域を移動することにより防止できる。

【0006】さらに、特開平2-201655号公報は、1台の計算機を構成する複数の中央処理装置を、仮想計算機の中央処理装置として割り付ける仮想計算機割り当て方式を開示している。

【0007】

【発明が解決しようとする課題】仮想計算機システムにおけるゲスト主記憶領域の割り当てには、以下の課題がある。

【0008】(1) 第1の課題は、ゲスト主記憶領域のサイズのトータルバランスに関する。

【0009】仮想計算機システムの運用では、仮想計算機の負荷変動が定期的に発生することが多い。例えば、

2台の仮想計算機が毎日運用される場合、第1の仮想計算機では、TSS (Time Sharing System) を運用しているために昼間の負荷が高く、第2の仮想計算機では、空き時間にバッチジョブを実行するために夜の負荷が高い、という使い方も多い。このような場合、負荷の高い方の仮想計算機に、より多くの主記憶領域を与えた方が、ページング処理等を削減でき、全体のスループットが向上する。

【0010】また、図1に示すように第1の計算機500に対して第2の計算機600上の仮想計算機VM-1がホットスタンバイ状態にある場合、第1の計算機500が正常に動作している間は、仮想計算機VM-1には、比較的少量のゲスト主記憶領域420-1を与え、バッチジョブやTSSを運用する仮想計算機VM-2のゲスト主記憶領域420-2を大きくしておいた方がよい。これは、仮想計算機VM-1では、第1の計算機500に障害が発生した場合に備えた処理のみを行なっているからである。第1の計算機500の障害によりシステム切り替え装置700が、第1の計算機500の処理を仮想計算機VM-1に切り替えた場合、この仮想計算機VM-1のゲスト主記憶領域420-1のサイズを大きくして、高い負荷に耐えられるようにすることが望ましい。

【0011】しかしながら、従来、仮想計算機のゲスト主記憶領域のサイズを変更するには、他の仮想計算機の主記憶領域の配置も変更しなければならないため、全ての仮想計算機を停止しなければならなかった。

【0012】(2) 第2の課題は、ゲスト主記憶領域の主記憶装置上での移動に関する。仮想計算機上の OS が入出力命令で実行を要求するチャネルプログラム (ゲストチャネルプログラム) は、ゲスト主記憶領域上の絶対アドレス (ゲスト絶対アドレス) で記述されている。このため、OS 発行の入出力命令の仮想計算機モニタ (VMM) によるシミュレーションでは、データアドレスを主記憶装置での絶対アドレス (ホスト絶対アドレス) に変換したチャネルプログラム (ホストチャネルプログラム) を作成し、次に、このホストチャネルプログラムを実行する。従って、ゲスト主記憶領域の移動前に実行を開始したホストチャネルプログラムのデータアドレスは、全て、移動前のゲスト主記憶領域を示している。このため、特開平3-211629号公報が開示する方法によって、ゲスト主記憶領域を移動後、仮想計算機モニタによりホストチャネルプログラムの実行を再開すると、移動前の領域のデータのみが更新され、移動後のゲスト主記憶領域のデータは更新されない。従って、データの不一致が発生してしまう。

【0013】(3) 第3の課題は、ページ固定 VM からページ可能 VM への変更に関する。新たな仮想計算機の生成等により主記憶領域が必要な場合、ページ固定 VM をページ可能 VM として、空き領域を確保することが望まし

い。この際、ゲスト主記憶領域のサイズ変更、移動等を伴うので、上記と同様の課題がある。

【0014】(4) 第4の課題は、ページ可能VMからページ固定VMへの変更に関する。主記憶装置上に十分な空き領域がある場合、性能を向上させるためにページ可能VMをページ固定VMとすることが望ましい。この場合にも、通常、ゲスト主記憶領域のサイズ変更、移動等を伴うので上記と同様の課題がある。

【0015】(5) 第5の課題は、仮想計算機への中央処理装置の割り当てに関する。仮想計算機の負荷変動が定期的に発生する場合や、ホットスタンバイ状態であった仮想計算機にシステムが切り替わった場合、高い負荷の仮想計算機に対して優先的に、中央処理装置を与えるべきである。しかしながら、特開平2-201655号公報が示すように、従来、仮想計算機への中央処理装置の割り当ては、仮想計算機の生成時に固定的に決められていた。

【0016】本発明の目的は、各仮想計算機の負荷変動に応じた計算機資源の再配分を可能とする、仮想計算機システムにおける計算機資源の動的割付け方法を提供することにある。

【0017】本発明の他の目的は、ゲスト主記憶領域のサイズを動的に変更可能な仮想計算機システムを提供することにある。

【0018】本発明のさらに他の目的は、チャネル等の非同期処理装置によるゲスト主記憶領域へのアクセスを停止することなく、ゲスト主記憶領域を主記憶装置上で移動可能な、仮想計算機システムにおける計算機資源の動的割付け方法を提供することにある。

【0019】また、本発明の別の目的は、仮想計算機への中央処理装置の割り当てを動的に変更可能な、仮想計算機システムにおける計算機資源の動的割付け方法を提供することにある。

【0020】

【課題を解決するための手段】本発明による仮想計算機システムにおける計算機資源の割付け方法は、中央処理装置と、主記憶装置と、前記中央処理装置からの指示に応じて当該中央処理装置と非同期に処理を行う非同期処理装置とを含む計算機上で複数のオペレーティング・システム(OS)が仮想計算機モニタの制御下で走行する仮想計算機システムにおいて、前記複数のOSのうち、第1のOSに対して、前記主記憶装置の第1の固定領域を割付けるとともに、第2のOSに対して前記第1の固定領域のアドレス高位側に隣接する第2の固定領域を割り付けるステップと、特定のイベントが発生した時に、前記第2のOSに割り当てられた第2の固定領域のサイズを縮小するステップと、該サイズを縮小した第2の固定領域を前記主記憶装置上でアドレス高位側へ移動することにより、前記第1の固定領域のアドレス高位側に空き領域を生成するステップと、前記第1のOSの主記憶

領域のサイズを拡大するステップとを有するものである。

【0021】前記空き領域を生成するステップは、好ましくは、前記第2のOSの主記憶領域を前記仮想計算機モニタが生成した仮想空間とし、該仮想空間を前記縮小された第2の固定領域にマッピングするステップと、前記仮想空間のマッピング先を前記主記憶装置の移動先の領域に変更するステップと、該移動先の領域を前記第2のOSの固定領域として割り付けるステップを有する。

【0022】また、前記第2のOSの主記憶領域の移動中は、前記第2のOSが発行した前記非同期処理装置の使用開始命令を、前記仮想計算機モニタがシミュレーションするステップと、該シミュレーション時に前記非同期処理装置が前記第2のOSの主記憶領域のうちどの領域をアクセスするかを解析するステップと、該アクセス対象となる領域は、前記非同期処理装置によるアクセス終了後、移動するステップとを有することが望ましい。

【0023】

【作用】

(1) 各仮想計算機の生成時等に、特定のイベントの発生前と発生後における、この仮想計算機の計算機資源を構成テーブルに定義しておく。そして、特定のイベントが発生した時に、第2のOS(仮想計算機)の主記憶領域(ゲスト主記憶領域)のサイズを構成テーブルの定義に従って縮小し、各仮想計算機のゲスト主記憶領域を主記憶装置上で移動することにより、第1の仮想計算機のゲスト主記憶領域に隣接するアドレス高位の領域を空き領域にする。その後、第1の仮想計算機のゲスト主記憶領域のサイズを構成テーブルの定義に従って拡大する。

【0024】また、このとき構成テーブルの定義に従って仮想計算機への中央処理装置の割り当てを変更することもできる。たとえば、3台の仮想計算機が2台の中央処理装置を共用していた状態から、第1と第2の仮想計算機が第1の中央処理装置を共用し、第1の仮想計算機が第2の中央処理装置を専有する状態に移移する。

【0025】ここで、特定のイベントとは、特定の時間になったというイベントや、仮想計算機が走行する計算機と異なる第2の計算機に障害が発生し、この第2の計算機の処理を第1の仮想計算機上のOSが引き継いだというイベントを示す。

【0026】以上のようにして、ゲスト主記憶領域のサイズを動的に変更可能な仮想計算機システムを提供できる。

【0027】また、仮想計算機への中央処理装置の割り当てを動的に変更可能な仮想計算機システムを提供できる。

【0028】(2) ページ固定VMのゲスト主記憶領域を、主記憶装置上で移動するとき、ゲスト主記憶領域と主記憶装置上の連続領域との対応関係を仮想計算機モニタ(VMM)が生成した仮想空間(ホスト仮想空間)に

定義し、ホスト仮想空間のマッピング先を移動先領域に変更し、移動後の主記憶装置上の連続領域をOSのゲスト主記憶領域として割り付ける。また、ゲスト主記憶領域の移動中は、OSによる非同期処理装置の使用開始命令は、仮想計算機モニタ（VMM）がシミュレーションし、非同期処理装置によるゲスト主記憶領域のアクセス対象領域は、そのアクセス完了後に、移動する。この際、ページ可能VMの利用により、ゲスト主記憶領域の移動中には、ホスト仮想空間に対応する主記憶領域はページ単位に不連続となることが許容される。このようにして、非同期処理装置によるゲスト主記憶領域のアクセス対象外の領域から順次、領域を移動する。このため、領域移動後に、非同期処理装置が誤って移動前の領域にアクセスすることはない。

【0029】以上のようにして、チャネル等の非同期処理装置によるゲスト主記憶領域へのアクセスを停止することなく、ゲスト主記憶領域を主記憶装置上で移動させることができる。

【0030】(3) ページ固定VMのゲスト主記憶領域とこれが常駐する主記憶装置上の連続領域の対応関係をホスト仮想空間に定義し、ホスト仮想空間の定義情報をアドレス変換に使用してOSを走行させる。また、ホスト仮想空間のマッピング先の領域をページアウトすることにより、ホスト仮想空間のマッピング先の領域サイズをゲスト主記憶領域サイズよりも小さくする。このようにして、動的にページ固定VMをページ可能VMに変更できる。

【0031】(4) ページ可能VMのゲスト主記憶領域と主記憶装置との対応関係を保持するホスト仮想空間のマッピング先の領域がページアウトされている場合、この領域を主記憶装置上にページインする。次に、ホスト仮想空間のマッピング先として、主記憶装置の連続領域を対応付ける。そして、この連続領域の主記憶装置における配置情報を使用してOSを走行させる。このようにして、動的にページ可能VMをページ固定VMに変更できる。

【0032】

【実施例】以下、本発明のゲスト主記憶領域の移動方法に関する第1の実施例と、計算機資源の割り付け変更方法に関する第2の実施例を、図を用いて詳細に説明する。

【0033】1. ゲスト主記憶領域の移動方法

図2に、本実施例の仮想計算機システム（VMS）の構成図を示す。図2において、100-1および100-2は中央処理装置、200は中央処理装置100-k（k=1, 2）とは非同期的に処理を行なう非同期処理装置、300は主記憶装置400に対する参照及び更新を制御する記憶制御装置である。

【0034】非同期処理装置200は、計算機が通常の汎用大型計算機の場合には、チャネルである。チャネル

は、中央処理装置100-k（k=1, 2）の動作と非同期に、入出力装置と主記憶装置400の間でデータ転送を行なう。また、計算機がスーパーコンピュータの場合には、非同期処理装置200は、ベクトルプロセッサである。ベクトルプロセッサは、中央処理装置100-k（k=1, 2）の動作と非同期に、主記憶装置400上の命令列（いわゆる、ベクトル命令列）を実行する。本実施例では、初期状態として3台のページ固定VMが走行しており、仮想計算機モニタ（VMM）410の領域を除く主記憶装置400の記憶領域は、仮想計算機VM-1のゲスト主記憶領域420-1、仮想計算機VM-2のゲスト主記憶領域420-2、仮想計算機VM-3のゲスト主記憶領域420-3として各仮想計算機に割り付けられているものとする。

【0035】次に、図3を用いて仮想計算機モニタ410が、各仮想計算機の走行を制御するために作成する制御ブロック430-i（i=1, 2, 3）について説明する。制御ブロック430-iは、各仮想計算機VM-i（i=1, 2, 3）に対応して存在し、仮想計算機VM-i上のOSが中央処理装置100-k（k=1, 2）において走行中の各種制御情報を与える。

【0036】図3において、直接実行ビット431が1のとき、仮想計算機上のOSが発行した非同期処理装置200の起動開始命令は直接実行され、0のとき、命令の実行は抑止されて、仮想計算機モニタ410を介して実行される。ページ固定VM識別ビット432が1であるとき仮想計算機は、ページ固定VMであることを示し、0のときページ可能VMであることを示す。ページ固定VM識別ビット432が1であるとき、起点アドレスフィールド433および終点アドレスフィールド434は、ゲスト主記憶領域の主記憶装置400における起点アドレスおよび終点アドレスを保持する。また、ページ固定VM識別子435が0であるとき、ホスト第1制御レジスタ435は、この仮想計算機のゲスト主記憶領域が常駐するホスト仮想空間を定義するホストセグメントテーブル（ホストST）の先頭アドレスを保持する。

【0037】次に、図4を用いて仮想計算機システム（VMS）におけるアドレス変換について説明する。

【0038】仮想計算機VM-2上のOSが生成した仮想空間（ゲスト仮想空間）上の論理アドレス（ゲスト論理アドレス）4000は、OSがゲスト主記憶領域420-2上に作成したゲストセグメントテーブル（ゲストST）およびゲストページテーブル（ゲストPT）によるアドレス変換4100の結果、ゲスト絶対アドレス4200に変換される。そして、仮想計算機がページ固定VMのとき、即ち、この仮想計算機の制御ブロック430のページ固定VM識別ビット432が1のとき、ページ固定VM用アドレス変換回路4300による変換結果がセレクタ4500により選択され、ホスト絶対アドレス4600となる。一方、仮想計算機がページ可能V

Mのとき、即ち、この仮想計算機の制御ブロック430のページ固定VM識別ビット432が0のとき、ページ可能VM用アドレス変換回路4400による変換結果がセクタ4500により有効化され、ホスト絶対アドレス4600となる。

【0039】ここで、ページ固定VM用アドレス変換回路4300によるゲスト絶対アドレス4200からホスト絶対アドレス4600への変換は、前記特開昭60-24735号公報に開示されているように、起点アドレスフィールド433の値を加算器4310によりゲスト絶対アドレス4200に加算することにより行なわれる。ただし、加算の結果の値が、終点アドレスフィールド434の値を超える場合、比較器4320は、プログラム割込みをOSに報告するために割込み信号4330を発生させる。

【0040】一方、ページ可能VM用アドレス変換回路4400によるゲスト絶対アドレス4200からホスト絶対アドレス4600への変換は、前記特開昭57-212680号公報に開示されているように、次のように行なわれる。まず、ゲスト絶対アドレス4200は、仮想計算機モニタ410からみればホスト仮想アドレスである。このため、ホスト第1制御レジスタ435の値とホスト仮想アドレス4410のセグメントフィールド

(S)の値を加算器4420により加算してホストセグメントテーブル(ホストST)4430のエントリアドレスを得る。このエントリのページテーブルアドレスとホスト仮想アドレス4410のページフィールド(P)の値を加算器4440により加算してホストページテーブル(ホストPT)4450のエントリアドレスを得る。さらにこのエントリのページアドレスとホスト仮想アドレス4410のディスプレイメント(D)を結合することにより、ホスト絶対アドレス4600を得る。

【0041】次に、図5および図6を用いて仮想計算機の主記憶領域の主記憶装置400上での移動方法を説明する。図5はゲスト主記憶領域の移動方法のフローチャートであり、図6はゲスト主記憶領域の移動方法の説明図である。

【0042】図6(a)に示す初期状態において仮想計算機VM-3が走行を終了し、図6(b)に示すように、主記憶装置400の一部の領域が空き状態になったとする。そこで、以下に述べる方法により、仮想計算機VM-2のゲスト主記憶領域420-2を主記憶装置400上で移動する。

【0043】まず、仮想計算機モニタ410は、仮想計算機VM-2上のOSが発行した非同期処理装置200の起動開始命令をシミュレーションするために、仮想計算機VM-2の制御ブロック430-2内の直接実行ビット431を0に設定する(ステップ5000)。これにより、ゲスト主記憶領域420-2の移動中に、仮想計算機VM-2上のOSが非同期処理装置200の起動

開始命令を発行すると、命令の実行は中止されて、仮想計算機モニタ410に割り込まれる。

【0044】次に、仮想計算機モニタ410は、ページ固定VMである仮想計算機VM-2に対して、図6

(c)に示すように、ホスト仮想空間450を生成し、VMをページ可能VMに変更する(ステップ5010)。このホスト仮想空間450を定義するホストセグメントテーブル(ホストST: Host Segment Table)およびホストページテーブル(ホストPT: Host Page Table)は、ホスト仮想アドレスに図6(c)に示すページング領域の起点アドレスを加算した値が、仮想計算機VM-2のゲスト主記憶領域のホスト絶対アドレスとなるようにテーブル値を設定し、仮想計算機モニタ410内の領域に作成する。即ち、ホスト仮想空間450の第jページが図6(c)に示すページング領域の第jページに対応付けられるように、仮想計算機モニタ410は、ホストSTおよびホストPTを作成する。

【0045】次に、仮想計算機モニタ410は、マッピング先の領域サイズを小さくする必要があるかを判定する(ステップ5020)。この判定は、仮想計算機システム(VMS)のユーザのコマンド入力等により、マッピング先の領域サイズを小さくするよう指定があったか否かにより判定する。指定があった場合、マッピング先の領域サイズを指定されたサイズに変更し、残りのフィールドはページアウトする。これによりゲスト主記憶領域のサイズよりもホスト仮想空間450のマッピング先の領域サイズの方を小さくする(ステップ5030)。ただし、非同期処理装置200によるアクセス対象のゲスト主記憶領域は、ページアウトの対象外とする。

【0046】次に、走行させるOSが、非同期処理装置200を使用中か否かを判定する(5035)。使用中でなければ、当該OSのホスト仮想空間の全領域のマッピング先を移動先領域に変更して(5040)、ステップ5100へ移行する。この移動先領域の変更は、移動前の領域の内容を移動後の領域にコピーし、ホストSTおよびホストPTの変更により、移動前の領域にマッピングされていたホスト仮想空間450上の領域を移動後の領域にマッピングすることにより行なう。ステップ5035で使用中ならば、図3の仮想計算機VM-2の制御ブロック430-2のページ固定VM識別ビット432に0を設定し、ホスト第1制御レジスタフィールド435にホストセグメントテーブル(ホストST)のアドレスを設定することによりホスト仮想空間450の定義情報を指定して仮想計算機VM-2上のOSをディスプレイする(ステップ5050)。その後、OSの走行(ステップ5060)が開始される。このOSが走行中は、仮想計算機VM-2の制御ブロック430-2内のページ固定VM識別ビット432は0なので、ゲスト絶対アドレスからホスト絶対アドレスへの変換は、仮想計算機VM-2の制御ブロック430-2のホスト第1制

御レジスタフィールド435に指定されたホストセグメントテーブル（ホストST）と、このホストSTが指定するホストページテーブル（ホストPT）に従って、行なわれる。

【0047】OS走行開始後、仮想計算機VM-2上のOSが非同期処理装置200の起動開始命令を発行（ステップ5070）すると、VM-2の制御ブロック430-2の直接実行ビット431は0なので、この命令は、仮想計算機モニタ410に割り出される。この結果、仮想計算機モニタ410は、非同期処理装置200の起動開始命令のシミュレーションを行なう（ステップ5080）。このシミュレーションでは、仮想空間の現在のマッピング状態に従ってデータ転送の対象とするアドレスの変換を行い、非同期処理装置200に起動開始命令を改めて発行する。起動が成功すれば（5083）、ステップ5085へ進み、未だ非同期処理装置使用中等の理由により起動に失敗すれば、ステップ5050へ戻る。

【0048】その後、図6（d）に示すように、非同期処理装置200のデータ転送に関係のないホスト仮想空間450のマッピング先を移動先領域に変更する（ステップ5085）。非同期処理装置200のデータ転送に関係のあるゲスト主記憶領域420-2内の領域は、非同期処理装置200の起動開始命令に指定されている。このため、該当するページ以外のマッピング先を、徐々に、移動先の領域に変更していく。

【0049】このシミュレーション後、仮想計算機モニタ410は、仮想計算機VM-2のゲスト主記憶領域420-2の移動を完了したかを判定する（ステップ5090）。この判定は、ホスト仮想空間450のマッピング先が、全て移動先の領域となったか否かで判定できる。そして、図6（e）に示すように移動完了時、制御ブロック430-2内の起点アドレスフィールド433および終点アドレスフィールド434に移動後のゲスト主記憶領域420-2の起点アドレスおよび終点アドレスを設定することにより、ゲスト主記憶装置420-2のアドレス変換情報を再設定する（ステップ5100）。次に、図6（f）に示すように、仮想計算機VM-2の制御ブロック430-2内のページ固定VM識別ビット432に1を設定し、ページ固定VMとする（ステップ5110）。また、仮想計算機VM-2上のOSによる非同期処理装置200の起動開始命令を直接実行するために、仮想計算機VM-2の制御ブロック430-2内の直接実行ビット431を1に設定する（ステップ5120）。

【0050】そして、図3の仮想計算機VM-2の制御ブロック430-2を指定してOSをディスパッチする（ステップ5130）。この後、OSの走行（ステップ5140）が開始される。このOSが走行中は、ゲスト絶対アドレスからホスト絶対アドレスへの変換は、仮想

計算機VM-2の制御ブロック430-2内のページ固定VM識別ビット432は1であるから、ゲスト絶対アドレスに起点アドレスフィールド433の値を加算することにより行なわれる。さらに、終点アドレスフィールド434の値と、ホスト絶対アドレスを比較することによって、アドレスの上限チェックが、行なわれる。

【0051】走行開始後、仮想計算機VM-2上のOSが非同期処理装置200の起動開始命令を発行すると（ステップ5150）、VM-2の制御ブロック430-2の直接実行ビット431は1なので、この命令は、仮想計算機モニタ410に割り出されることがなく、命令処理装置100によって、直接実行される（ステップ5160）。

【0052】以上のようにして、本実施例によれば、中央処理装置とは非同期に処理を行なうチャネル等の非同期処理装置200によるゲスト主記憶領域へのアクセスを停止することなく、ゲスト主記憶領域を主記憶装置上で移動可能な仮想計算機システムを提供できる。また、動的に、ページ固定VMをページ可能VMに変更し、あるいは、ページ可能VMをページ固定VMに変更することができる。

【0053】II. 計算機資源の割り付け変更方法

図7、図8、および、図9を用いて計算機資源の割り付け変更方法を説明する。図7は計算機資源の割り付け変更方法のフローチャートであり、図8はゲスト主記憶領域のサイズの変更方法の説明図であり、図9は構成テーブル9000の説明図である。構成テーブル9000は、イベント発生前とイベント発生後において各仮想計算機に割り付ける計算機資源を示す。例えば、図9の構成テーブル9000は、図8（a）に示す初期状態（イベント発生前の状態）において、32メガバイトからなるゲスト主記憶領域420-1を有する仮想計算機VM-1と、128メガバイトからなるゲスト主記憶領域420-2を有する仮想計算機VM-2と、32メガバイトからなるゲスト主記憶領域420-1を有する仮想計算機VM-3の3台の仮想計算機が、中央処理装置100-1および100-2を共用して走行していることを表わしている。

【0054】この後、特定のイベントが発生すると仮想計算機モニタ410は、仮想計算機VM-2上のOSにゲスト主記憶領域420-2を構成テーブルに指定された32メガバイトに縮小するよう要求する（ステップ7000）。ここで、特定のイベントとは、特定の時間になったというイベントや、図1に示す第1の計算機500に障害が発生し、システム切り替え装置700によって、この第1の計算機500の処理を第2の計算機600の仮想計算機VM-1上のOSが引き継いだというイベントを示す。この結果、仮想計算機VM-2上のOSは、図8（b）に示すように、ゲスト主記憶領域420-2内の高位アドレスの領域をサテライン（示線状態

態)にし(ステップ7010)、仮想計算機VM-2から切り離す。

【0055】次に、仮想計算機モニタ410は、仮想計算機VM-2のゲスト主記憶領域420-2を、第1の実施例の場合と同様に、移動する(ステップ7020)。これによって、図8の(c)に示すように、仮想計算機VM-1のゲスト主記憶領域420-1に隣接した高位アドレスの領域を空き領域とすることができる。

【0056】また、仮想計算機モニタ410が、仮想計算機VM-1上のOSにゲスト主記憶領域420-1を構成テーブルに指定された128メガバイトに拡大するよう要求する(ステップ7030)。この結果、図8の(d)に示すように、仮想計算機VM-1上のOSが、ゲスト主記憶領域420-1に隣接した高位アドレスの領域をオンライン(接続状態)にし(ステップ7040)、仮想計算機VM-1に接続する。

【0057】さらに、仮想計算機モニタ410は、構成テーブル9000が示すように、1台の中央処理装置100-1を仮想計算機VM-1に専有させ、中央処理装置100-2を仮想計算機VM-2と仮想計算機VM-3で共有させて、仮想計算機上のOSを走行させる(ステップ7050)。

【0058】以上のようにして、本実施例によれば、ゲスト主記憶領域のサイズを動的に変更可能な仮想計算機システムを提供できる。また、仮想計算機への中央処理装置の割り付けを動的に変更可能な仮想計算機システムを提供できる。

【0059】

【発明の効果】本発明によれば、各仮想計算機のゲスト主記憶領域を主記憶装置上で移動して特定の仮想計算機のゲスト主記憶領域に隣接するアドレス高位の領域を空き領域にすることにより、ゲスト主記憶領域のサイズを動的に変更可能な仮想計算機システムを提供できる。

【0060】また、チャネル等の非同期処理装置によるゲスト主記憶領域のアクセス対象外の領域から順次、領域を移動することにより、非同期処理装置によるゲスト主記憶領域へのアクセスを停止することなく、ゲスト主記憶領域を主記憶装置上で移動可能な仮想計算機システムを提供できる。

ムを提供できる。

【0061】また、ページ固定VMのゲスト主記憶領域とこれが等値する主記憶装置上の連続領域の対応関係をホスト仮想空間に定義することにより、ページ固定VMをページ可能VMに変更可能な仮想計算機システムを提供できる。

【0062】また、ホスト仮想空間のマッピング先として、主記憶領域の連続領域を対応付けることにより、ページ可能VMをページ固定VMに変更可能な仮想計算機システムを提供できる。

【0063】さらに、イベントの発生を契機として各仮想計算機に構成テーブルに記載された計算機資源を割り付けることにより、仮想計算機への中央処理装置の割り付けを動的に変更可能な仮想計算機システムを提供できる。

【図面の簡単な説明】

【図1】本発明の適用対象を例示するためのホットスタンバイシステムの説明図。

【図2】本発明が適用される仮想計算機システムの説明図。

【図3】図2の仮想計算機の制御ブロックの説明図。

【図4】図2の仮想計算機システムにおけるアドレス変換の説明図。

【図5】本発明の第1の実施例におけるゲスト主記憶領域の移動方法のフローチャート。

【図6】本発明の第1の実施例におけるゲスト主記憶領域の移動方法の説明図。

【図7】本発明の第2の実施例における仮想計算機に対する計算機資源の割付け変更方法のフローチャート。

【図8】本発明の第2の実施例におけるゲスト主記憶領域のサイズの変更方法の説明図。

【図9】本発明の第2の実施例における構成テーブルの説明図。

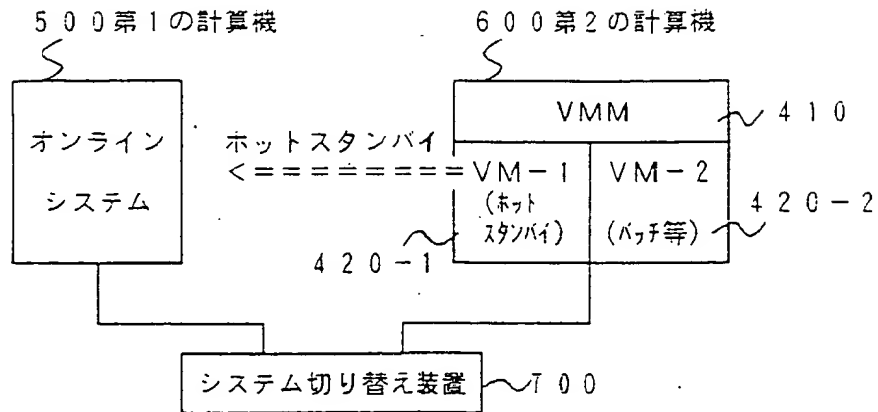
【符号の説明】

100：中央処理装置、200：非同期処理装置、300：記憶制御装置、400：主記憶装置、410：仮想計算機モニタ、420-i：仮想計算機VM-iのゲスト主記憶領域、9000：構成テーブル

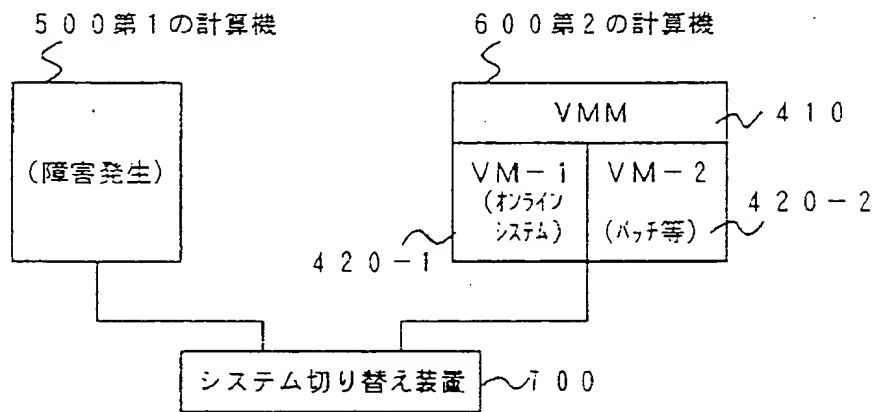
【図1】

ホットスタンバイシステムの説明図（図1）

(a) システム切り替え前

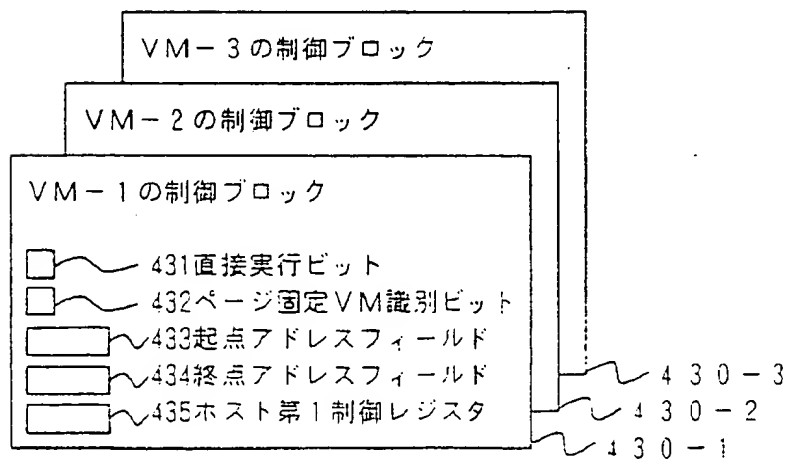


(b) システム切り替え後



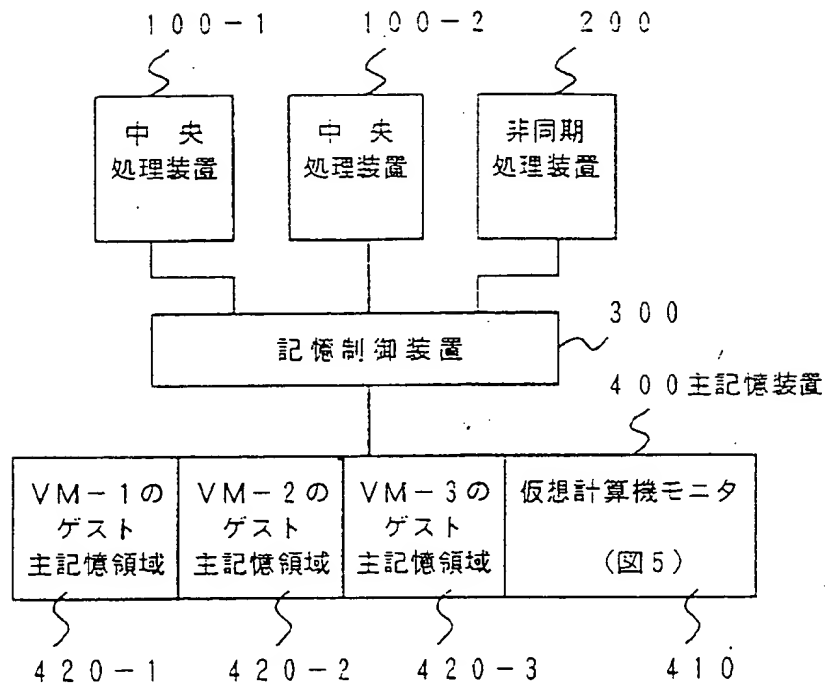
【図3】

仮想計算機（VM）の制御ブロックの説明図（図3）



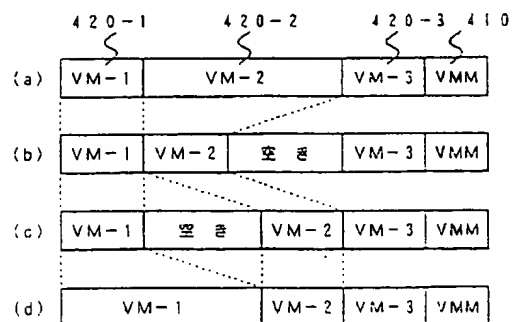
【図2】

仮想計算機システムの説明図（図2）



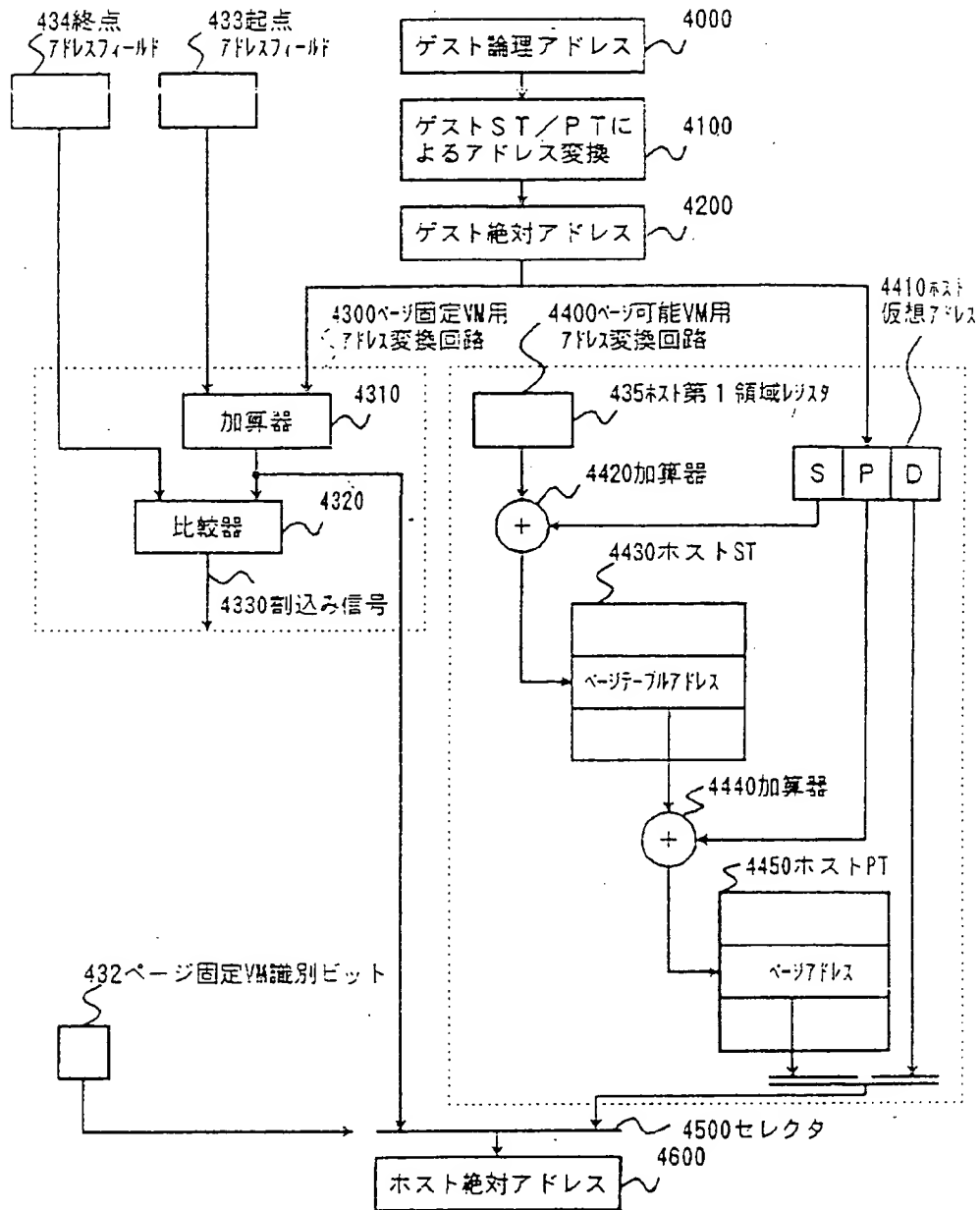
【図8】

ゲスト主記憶領域のサイズの変更方法の説明図（図8）



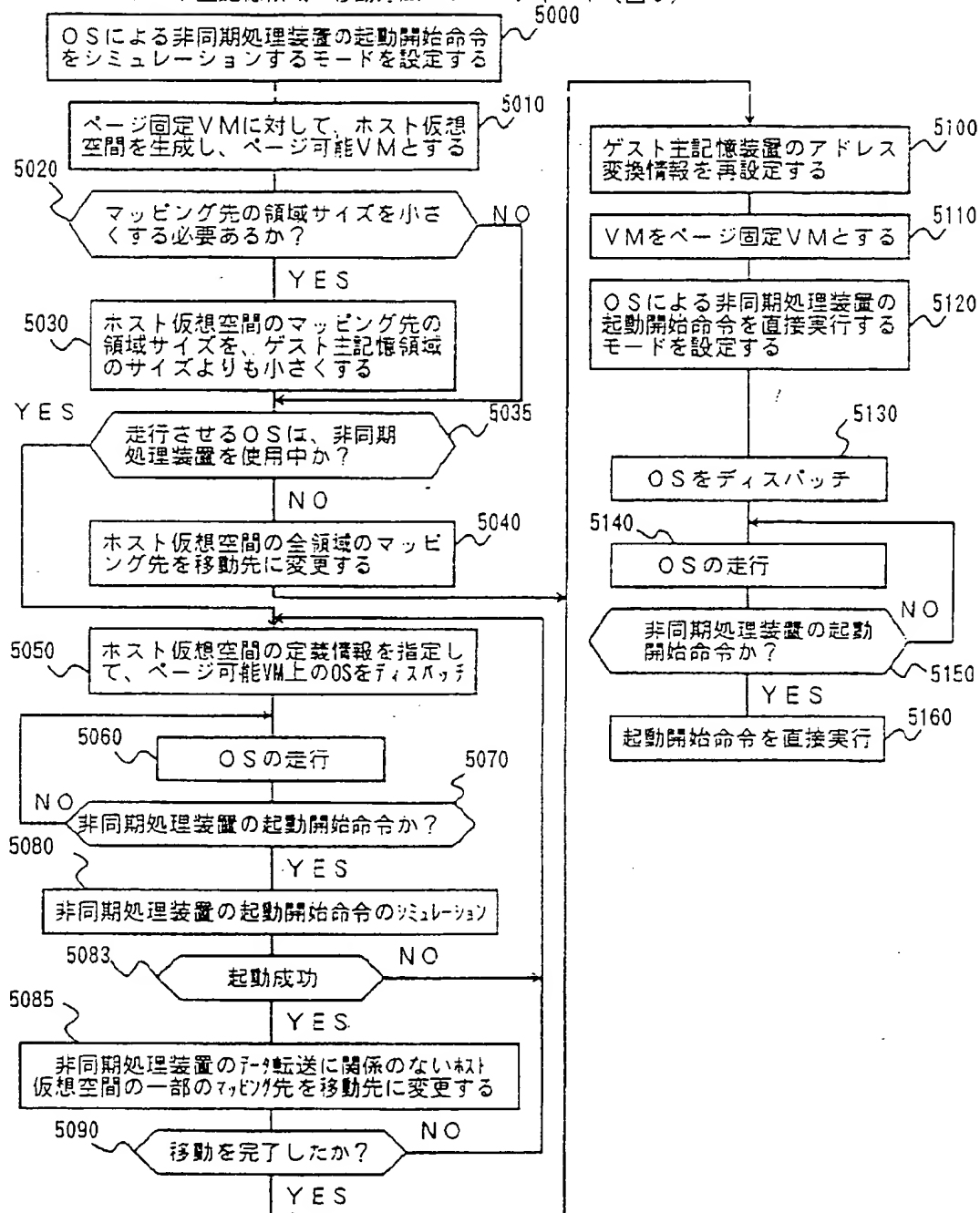
【図4】

仮想計算機システムにおけるアドレス変換の説明図（図4）



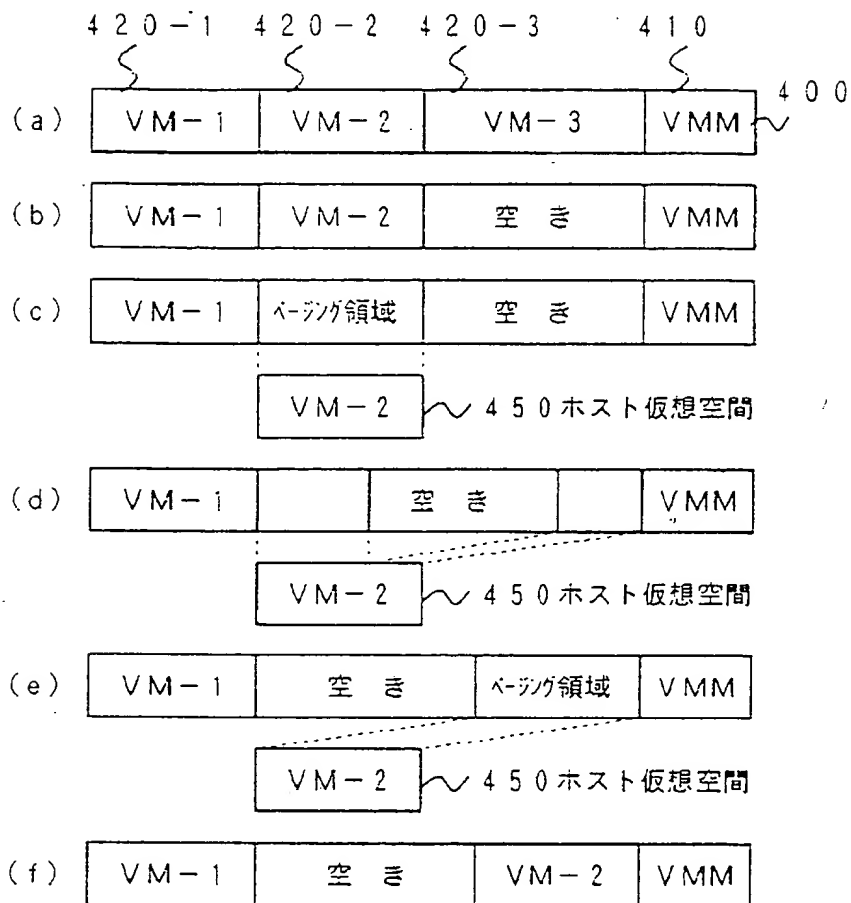
【図5】

ゲスト主記憶領域の移動方法のフローチャート（図5）



【図6】

ゲスト主記憶領域の移動方法の説明図（図6）



【図9】

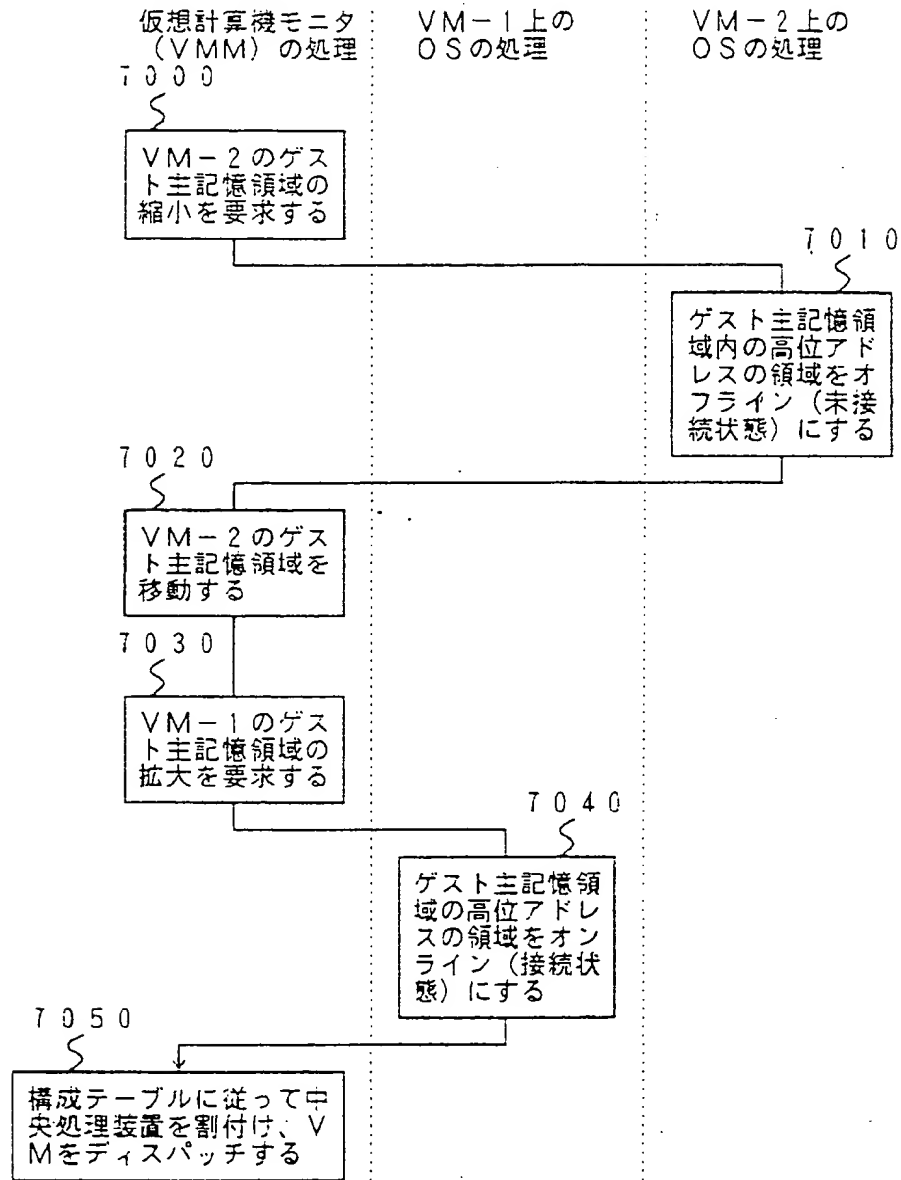
構成テーブルの説明図（図9）

9000構成テーブル

		VM-1	VM-2	VM-3
イベント発生前	ゲスト主記憶サイズ	32MB	128MB	32MB
	中央処理装置の割り当て	共用	共用	共用
イベント発生後	ゲスト主記憶サイズ	128MB	32MB	32MB
	中央処理装置の割り当て	専有	共用	共用

【図7】

計算機資源の割り当て変更方法のフローチャート（図7）



フロントページの続き

(72)発明者 山本 正

神奈川県横浜市戸塚区戸塚町5030番地 株
式会社日立製作所ソフトウェア開発本部内

(72)発明者 池ヶ谷 直子

神奈川県川崎市麻生区王禅寺1099 株式会
社日立製作所システム開発研究所内

[Brief Description of the Drawings]

[Drawing 1] The explanatory view of the hot standby system for illustrating the object for application of this invention.

[Drawing 2] The explanatory view of a virtual machine system in which this invention is applied.

[Drawing 3] The explanatory view of control block of the virtual machine of drawing 2.

[Drawing 4] The explanatory view of the address translation in the virtual machine system of drawing 2.

[Drawing 5] The flow chart of handling of the guest main storage area in the 1st example of this invention.

[Drawing 6] The explanatory view of handling of the guest main storage area in the 1st example of this invention.

[Drawing 7] The flow chart of the allocation modification approach of a calculating-machine resource for the virtual machine in the 2nd example of this invention.

[Drawing 8] The explanatory view of the modification approach of the size of the guest main storage area in the 2nd example of this invention.

[Drawing 9] The explanatory view of the configuration table in the 2nd example of this invention.

[Description of Notations]

100: A central processing unit, 200: asynchronous process equipment, a 300: memory control unit, 400: main storage, a 410: virtual-machine monitor, the guest main storage area of 420-i: virtual-machine VM-i, 9000 : configuration table

JPO translation of
Tanaka et al. (cited by Applicant IDS)

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to a suitable computing system to change allocation of the computer resource to OS dynamically about a computing system in the computing system two or more OS's run on one computer especially.

[0002]

[Description of the Prior Art] As a computing system which can run two or more operating systems (OS: Operating System) in an one calculating-machine top, there is a virtual machine system (VMS: Virtual Machine System). A virtual machine system enables transit of one OS which generates two or more virtual machines (VM: Virtual Machine) which are logical computers, and corresponds in each virtual machine (VM) on one computer. This virtual machine system is mainly used, in order to employ two or more different activity eye OS's on one calculating machine, or to perform system generation of OS of an upgrade product, employing OS of an earlier version or to perform the test of OS to juxtaposition on one calculating machine. In a virtual machine system, the virtual-machine monitor (VMM: Virtual Machine Monitor) which is the control program distributes a calculating-machine resource to each virtual machine (VM), and performs scheduling of a virtual machine (VM), and dispatch. Moreover, a virtual-machine monitor (VMM) performs simulation processing of the instruction which cannot carry out immediate execution, i.e., the instruction which needs inclusion of software, etc. by the computer among the instructions which OS on a virtual machine published.

[0003] A virtual machine has page possible VM and the page immobilization VM depending on how to give the main storage area (guest main storage area) of a virtual machine. Page possible VM stations a guest main storage area permanently like a JP,57-212680, A publication on the virtual space (host virtual space) which the virtual-machine monitor (VMM) generated. Therefore, the guest main storage area of page possible VM is set as the object of paging by the virtual-machine monitor (VMM). By this paging, size of the guest main storage area of page possible VM is made more greatly than the size of the paging area of the virtual-machine monitor (VMM) which is the mapping place of a host virtual space.

[0004] On the other hand, the page immobilization VM stations a guest main storage area permanently at the continuation field on main storage like a JP,60-24735, A publication. Therefore, the guest main storage area of the page immobilization VM is not set as the object of paging by the virtual-machine monitor (VMM). For this reason, the page immobilization VM has few overheads which address translation takes compared with page possible VM. However, only the continuation area size of the matched main storage can set up the size of the guest main storage area of the page immobilization VM.

[0005] Moreover, JP,3-211629, A suspends access to the guest main storage area by asynchronous process equipments which process to asynchronous with a central processing unit, such as a channel, and is indicating the virtual machine system which makes movable the guest main storage area of the page immobilization VM on main storage. According to this virtual machine system, fragmentation of the storage region of main storage can be prevented by moving in a guest main storage area.

[0006] Furthermore, JP,2-201655, A is indicating the virtual-machine quota method which assigns two or more central processing units which constitute one computer as a central processing unit of a virtual machine.

[0007]

[Problem(s) to be Solved by the Invention] There are the following technical problems in assignment of the guest main storage area in a virtual machine system.

[0008] (1) The 1st technical problem is related with the total balance of the size of a guest main storage area.

[0009] In employment of a virtual machine system, the load effect of a virtual machine occurs periodically in many cases. For example, when two sets of virtual machines are employed all day, with the 1st virtual machine, since TSS (Time Sharing System) is employed, the load of day ranges is expensive, and in order to perform a batch job at idle time, with the 2nd virtual machine, the load of night uses often [of being high]. In such a case, the direction which gave more main storage areas to the virtual machine of the one where a load is more expensive can reduce paging processings etc., and

the whole throughput improves.

[0010] Moreover, it is better to enlarge the guest main storage area 420-2 of virtual-machine VM-2 which give the comparatively little guest main storage area 420-1 to virtual-machine VM-1, and employ a batch job and TSS, while the 1st calculating machine 500 is operating normally, when virtual-machine VM-1 on the 2nd calculating machine 600 is in a hot standby condition to the 1st calculating machine 500, as shown in drawing 1. This is because only processing which it had when a failure occurred to the 1st computer 500 is performed in virtual-machine VM-1. When the system transfer device 700 changes processing of the 1st computer 500 to virtual-machine VM-1 according to the failure of the 1st computer 500, it is desirable to enlarge size of this guest main storage area 420-1 of virtual-machine VM-1, and to enable it to bear a high load.

allocation
of
storage
to
different
vms

[0011] However, in order to change the size of the guest main storage area of a virtual machine conventionally and also to have to change arrangement of the main storage area of other virtual machines, all the virtual machines had to be suspended.

[0012] (2) The 2nd technical problem is related with the migration on the main storage of a guest main storage area. The channel program (guest channel program) as which OS on a virtual machine requires activation by input/output instruction is described by the absolute address on a guest main storage area (guest absolute address). For this reason, in the simulation by the virtual-machine monitor (VMM) of the input/output instruction of OS issuance, the channel program (host channel program) which changed the data address into the absolute address (host absolute address) in main storage is created, next this host channel program is performed. Therefore, all the data addresses of the host channel program which started activation before migration of a guest main storage area show the guest main storage area before migration. For this reason, if the guest main storage area migration-back is resumed by the approach which JP,3-211629,A indicates and activation of a host channel program is resumed with a virtual-machine monitor, only the data of the field before migration will be updated and the data of the guest main storage area after migration will not be updated. Therefore, the inequality of data will occur.

[0013] (3) The 3rd technical problem is related with modification to page possible VM from the page immobilization VM. It is desirable to make a free area by setting page immobilization VM to page possible VM by generation of a new virtual machine etc., when a main storage area is required. Under the present circumstances, since it is accompanied by size change of a guest main storage area, migration, etc., the same technical problem as the above occurs.

[0014] (4) The 4th technical problem is related with modification to the page immobilization VM from page possible VM. When there is free area sufficient on main storage, in order to raise the engine performance, it is desirable to consider page possible VM as the page immobilization VM. Also in this case, since it is accompanied by size change of a guest main storage area, migration, etc., the same technical problem as the above usually occurs.

[0015] (5) The 5th technical problem is related with assignment of the central processing unit to a virtual machine. When the load effect of a virtual machine occurs periodically, or when a system changes to the virtual machine which was in the hot standby condition, the central processing unit should be preferentially given to the virtual machine of a high load. However, as JP,2-201655,A showed, assignment of the central processing unit to a virtual machine was conventionally decided to be the generate time of a virtual machine fixed.

[0016] The object of this invention is to offer the dynamic allocation approach of the computer resource in a virtual machine system which makes possible reallocation of the computer resource according to the load effect of each virtual machine.

[0017] Other objects of this invention are about the size of a guest main storage area to offer the virtual machine system which can be changed dynamically.

[0018] The object of further others of this invention is to offer the dynamic allocation approach of a computer resource [in / for a guest main storage area / a movable virtual machine system] on main storage, without suspending access to the guest main storage area by asynchronous process equipments, such as a channel.

[0019] Moreover, another object of this invention is to offer the dynamic allocation approach of a

computer resource [in / for assignment of the central processing unit to a virtual machine / the virtual machine system which can be changed dynamically].

[0020]

[Means for Solving the Problem] An approach to assign the computer resource in the virtual machine system by this invention In the virtual machine system two or more operating systems (OS) in the calculating-machine top containing the central processing unit concerned and the asynchronous process equipment which processes to asynchronous run under control of a virtual-machine monitor according to the directions from a central processing unit, main storage, and said central processing unit While assigning the 1st fixed area of said main storage to the 1st OS among said two or more OS's The step which assigns the 2nd fixed area which adjoins the address high order side of said 1st fixed area to the 2nd OS, By moving the step which reduces the size of the 2nd fixed area assigned to said 2nd OS, and the 2nd fixed area which reduced this size to an address high order side on said main storage, when a specific event occurs It has the step which generates a free area to the address high order side of said 1st fixed area, and the step to which the size of the main storage area of said 1st OS is expanded.

[0021] The step which generates said free area is preferably made into the virtual space where said virtual-machine monitor generated the main storage area of said 2nd OS, and it has the step which maps this virtual space in said 2nd reduced fixed area, the step which changes the mapping place of said virtual space into the field of the migration place of said main storage, and the step which assigns the field of this migration place as a fixed area of said 2nd OS.

[0022] Moreover, during migration of the main storage area of said 2nd OS The step to which said virtual-machine monitor carries out simulation of the beginning-of-using instruction of said asynchronous process equipment which said 2nd OS published, As for the step which analyzes which field said asynchronous process equipment accesses among the main storage areas of said 2nd OS at the time of this simulation, and the field used as this object for access, it is desirable to have the step which moves after the access termination by said asynchronous process equipment.

[0023]

[Function]

(1) Define the calculating-machine resource of this virtual machine generating before of a specific event, and after generating as the configuration table at the generate time of each virtual machine etc. And when a specific event occurs, the field of an address high order contiguous to the guest main storage area of the 1st virtual machine is made into a free area by reducing the size of the main storage area (guest main storage area) of the 2nd OS (virtual machine) according to the definition of a configuration table, and moving on main storage in the guest main storage area of each virtual machine. Then, the size of the guest main storage area of the 1st virtual machine is expanded according to the definition of a configuration table.

[0024] Moreover, according to the definition of a configuration table, assignment of the central processing unit to a virtual machine can also be changed at this time. For example, from the condition that three sets of virtual machines were sharing two sets of central processing units, the 1st and the 2nd virtual machine share the 1st central processing unit, and the 1st virtual machine changes in the condition of having the 2nd central processing unit chiefly.

[0025] Here, a failure occurs to the event of having become specific time amount, the computer a virtual machine runs, and the 2nd different computer, and a specific event shows the event that OS on the 1st virtual machine succeeded processing of this 2nd computer.

[0026] The virtual machine system which can be changed dynamically can be offered for the size of a guest main storage area as mentioned above.

[0027] Moreover, the virtual machine system which can be changed dynamically can be offered for assignment of the central processing unit to a virtual machine.

[0028] (2) When moving on main storage in the guest main storage area of the page immobilization VM, define it as the virtual space (host virtual space) where the virtual-machine monitor (VMM) generated the response relation between a guest main storage area and the continuation field on main storage, change the mapping place of a host virtual space into a migration place field, and assign the continuation

field on the main storage after migration as a guest main storage area of OS. Moreover, a virtual-machine monitor (VMM) carries out simulation of the beginning-of-using instruction of asynchronous process equipment according to OS in under migration of a guest main storage area, and it moves after the completion of access in the access object domain of the guest main storage area by asynchronous process equipment. Under the present circumstances, during migration of a guest main storage area, becoming discontinuous [the main storage area corresponding to a host virtual space] per page is permitted by utilization of page possible VM. Thus, it moves one by one in a field from the field of the outside for [of the guest main storage area by asynchronous process equipment] access. For this reason, asynchronous process equipment does not access to the field before migration accidentally after a domain migration.

{ [0029] A guest main storage area can be moved on main storage, without suspending access to the guest main storage area according to asynchronous process equipments, such as a channel, as mentioned above.

[0030] (3) Define the response relation between the guest main storage area of the page immobilization VM, and the continuation field on the main storage with which this resides permanently as a host virtual space, and run address translation OS using the definition information on a host virtual space. Moreover, area size of the mapping place of a host virtual space is made smaller than guest main storage area size by carrying out page-out of the field of the mapping place of a host virtual space. Thus, the page immobilization VM can be dynamically changed into page possible VM.

[0031] (4) When page-out of the field of the mapping place holding the response relation between the guest main storage area of page possible VM and main storage of a host virtual space is carried out, carry out page-in of this field on main storage. Next, the continuation field of main storage is matched as a mapping place of a host virtual space. And it is made to run OS using the arrangement information in the main storage of this continuation field. Thus, page possible VM can be dynamically changed into the page immobilization VM.

[0032]

[Example] Hereafter, the 1st example about handling of the guest main storage area of this invention and the 2nd example about the allotment modification approach of a computer resource are explained to a detail using drawing.

[0033] I. The block diagram of the virtual machine system (VMS) of this example is shown in handling drawing 2 of a guest main storage area. As for 100-1 and 100-2, in drawing 2, a central processing unit and 200 are [the asynchronous process equipment which processes in / central processing unit 100-k (1 k= 2) / asynchronous, and 300] memory control units which control the reference and updating to main storage 400.

[0034] Asynchronous process equipment 200 is a channel when a calculating machine is the usual general purpose large-sized calculating machine. A channel performs data transfer to asynchronous [of central processing unit 100-k (1 k= 2) / actuation and asynchronous] between an I/O device and main storage 400. Moreover, when a calculating machine is a supercomputer, asynchronous process equipment 200 is a vector processor. A vector processor performs the instruction train on main storage 400 (the so-called vector instruction train) to asynchronous [of central processing unit 100-k (1 k= 2) / actuation and asynchronous]. In this example, three sets of the page immobilization VM shall run as an initial state, and the storage region of the main storage 400 except the field of the virtual-machine monitor (VMM) 410 shall be assigned to each virtual machine as the guest main storage area 420-1 of virtual-machine VM-1, the guest main storage area 420-2 of virtual-machine VM-2, and a guest main storage area 420-3 of virtual-machine VM-3.

[0035] Next, control-block 430-i (2 i= 1, 3) created in order that the virtual-machine monitor 410 may control transit of each virtual machine using drawing 3 is explained. Control-block 430-i exists corresponding to each virtual-machine VM-i (2 i= 1, 3), and OS on virtual-machine VM-i gives the various control information under transit in central processing unit 100-k (1 k= 2).

[0036] In drawing 3, when the immediate execution bit 431 is 1, immediate execution of the starting initiation instruction of the asynchronous process equipment 200 which OS on a virtual machine

published is carried out, and at the time of 0, activation of an instruction is inhibited and is performed through the virtual-machine monitor 410. When the page fixed VM discernment bit 432 is 1, it is shown that a virtual machine is the page immobilization VM, and it is shown that it is page possible VM at the time of 0. When the page fixed VM discernment bit 432 is 1, origin address field 433 and the terminal point address field 434 hold the origin address and the destination address in main storage 400 of a guest main storage area. Moreover, when the page fixed VM identifier 435 is 0, the 1st control register 435 of a host holds the start address of a host segment table (host ST) which defines the host virtual space in which the guest main storage area of this virtual machine resides permanently.

[0037] Next, the address translation in a virtual machine system (VMS) is explained using drawing 4.

[0038] The logical address 4000 (guest logical address) on the virtual space (guest virtual space) which OS on virtual-machine VM-2 generated is changed into the guest absolute address 4200 as a result of the address translation 4100 by the guest segment table (guest ST) and guest page table (guest PT) which OS created on the guest main storage area 420-2. And when a virtual machine is the page immobilization VM (i.e., when the page fixed VM discernment bit 432 of the control block 430 of this virtual machine is 1), a selector 4500 comes into effect and the conversion result by the address translation circuit 4300 for page fixed VM serves as the host absolute address 4600. On the other hand, when a virtual machine is page possible VM (i.e., when the page fixed VM discernment bit 432 of the control block 430 of this virtual machine is 0), a selector 4500 comes into effect and the conversion result by the address translation circuit 4400 for page possible VM serves as the host absolute address 4600.

[0039] Here, conversion to the host absolute address 4600 from the guest absolute address 4200 by the address translation circuit 4300 for page fixed VM is performed by adding the value of origin address field 433 to the guest absolute address 4200 with an adder 4310 as indicated by said JP,60-24735,A. However, when the value as a result of addition exceeds the value of the terminal point address field 434, a comparator 4320 generates an interrupt signal 4330, in order to report program interruption to OS.

[0040] On the other hand, conversion to the host absolute address 4600 from the guest absolute address 4200 by the address translation circuit 4400 for page possible VM is performed as follows as indicated by said JP,57-212680,A. First, the guest absolute address 4200 will be the host virtual address, if it sees from the virtual-machine monitor 410. For this reason, the value of the 1st control register 435 of a host and the value of the segment field (S) of the host virtual address 4410 are added with an adder 4420, and the entry address of the host segment table (host ST) 4430 is obtained. The value of the page field (P) of the page table address of this entry and the host virtual address 4410 is added with an adder 4440, and the entry address of the host page table (host PT) 4450 is obtained. The host absolute address 4600 is obtained by furthermore combining the displacement (D) of the page address of this entry, and the host virtual address 4410.

[0041] Next, handling on the main storage 400 of the main storage area of a virtual machine is explained using drawing 5 and drawing 6. Drawing 5 is the flow chart of handling of a guest main storage area, and drawing 6 is the explanatory view of handling of a guest main storage area.

[0042] In the initial state shown in drawing 6 (a), virtual-machine VM-3 end transit, and as shown in drawing 6 (b), suppose that some fields of main storage 400 would be in idle status. Then, it moves on main storage 400 by the approach described below in the guest main storage area 420-2 of virtual-machine VM-2.

[0043] First, in order that the virtual-machine monitor 410 may carry out simulation of the starting initiation instruction of the asynchronous process equipment 200 which OS on virtual-machine VM-2 published, the immediate execution bit 431 in virtual-machine VM-control block 430-2 of 2 is set as 0 (step 5000). Thereby, if OS on virtual-machine VM-2 publishes the starting initiation instruction of asynchronous process equipment 200 during migration of the guest main storage area 420-2, activation of an instruction will be inhibited and will be deduced by the virtual-machine monitor 410.

[0044] Next, to virtual-machine VM-2 which are the page immobilization VM, as shown in drawing 6 (c), the virtual-machine monitor 410 generates the host virtual space 450, and changes VM into page

possible VM (step 5010). The value which added the origin address of a paging area shown in drawing 6 (c) to the host virtual address sets up a table value so that it may become the host absolute address of the guest main storage area of virtual-machine VM-2, and it creates the host segment table (host ST:Host Segment Table) and host page table (host PT:Host Page Table) which define this host virtual space 450 in the field in the virtual-machine monitor 410. namely, the j-th page of a paging area which the j-th page of the host virtual space 450 shows to drawing 6 (c) -- matching ** -- the virtual-machine monitor 410 creates Host ST and Host PT like.

[0045] Next, it judges whether the virtual-machine monitor 410 needs to make area size of a mapping place small (step 5020). This judgment is judged by making area size of a mapping place small, and whether there was any assignment by the command input of the user of a virtual machine system (VMS) etc. When there is assignment, it changes into the size which had the area size of a mapping place specified, and page-out of the remaining fields is carried out. Thereby, the area size of the mapping place of the host virtual space 450 is made smaller than the size of a guest main storage area (step 5030). However, the guest main storage area for [by asynchronous process equipment 200] access is made into the outside of the object of page-out.

[0046] Next, OS it is made to run judges whether it is under [activity] ***** for asynchronous process equipment 200 (5035). If it is not [be / it] under activity, the mapping place of all the fields of the host virtual space of the OS concerned will be changed into a migration place field (5040), and it will shift to step 5100. Modification of this migration place field is copied to the field after moving the content of the field before migration, and is performed by mapping to the field after moving in the field on the host virtual space 450 mapped by modification of Host ST and Host PT to the field before migration. If it becomes during an activity at step 5035, by setting 0 as the page fixed VM discernment bit 432 of the control block 430-2 of virtual-machine VM-2 of drawing 3, and setting the address of a host segment table (host ST) as the 1st control register field 435 of a host, the definition information on the host virtual space 450 will be specified, and OS on virtual-machine VM-2 will be dispatched (step 5050). Then, transit (step 5060) of OS is started. Since the page fixed VM discernment bit 432 in virtual-machine VM-control block 430-2 of 2 is 0 while this OS runs, conversion to the host absolute address from the guest absolute address is performed according to the host segment table (host ST) specified as the 1st control register field 435 of a host of the control block 430-2 of virtual-machine VM-2, and the host page table (host PT) specified by this host ST.

[0047] After OS transit initiation, since the immediate execution bit 431 of the control block 430-2 of VM-2 is 0 when OS on virtual-machine VM-2 publishes the starting initiation instruction of asynchronous process equipment 200 (step 5070), this instruction is deduced by the virtual-machine monitor 410. Consequently, the virtual-machine monitor 410 performs simulation of a starting initiation instruction of asynchronous process equipment 200 (step 5080). In this simulation, the address made into the object of data transfer according to the current mapping condition of a virtual space is changed, and a starting initiation instruction is anew published to asynchronous process equipment 200. If starting is successful (5083), it will progress to step 5085, and if starting still goes wrong for the reason of the asynchronous process equipment activity middle class, it will return to step 5050.

[0048] Then, as shown in drawing 6 (d), the mapping place of the host virtual space 450 which is unrelated to the data transfer of asynchronous process equipment 200 is changed into a migration place field (step 5085). The field in the guest main storage area 420-2 which is related to the data transfer of asynchronous process equipment 200 is specified as the starting initiation instruction of asynchronous process equipment 200. For this reason, mapping places other than the corresponding page are gradually changed into the field of a migration place.

[0049] It judges whether the virtual-machine monitor 410 completed migration of the guest main storage area 420-2 of virtual-machine VM-2 after this simulation (step 5090). All the mapping places of the host virtual space 450 can judge this judgment by whether it became the field of a migration place. And as shown in drawing 6 (e), the address translation information on the guest main storage 420-2 is reset by setting up the origin address and the destination address of the guest main storage area 420-2 after moving to the origin address field 433 and the terminal point address field 434 in control-block 430-2 at

the time of the completion of migration (step 5100). Next, as shown in drawing 6 (f), 1 is set as the page fixed VM discernment bit 432 in control-block 430-2 of virtual-machine VM-2, and it considers as the page immobilization VM (step 5110). Moreover, in order to carry out immediate execution of the starting initiation instruction of the asynchronous process equipment 200 by OS on virtual-machine VM-2, the immediate execution bit 431 in virtual-machine VM-control block 430-2 of 2 is set as 1 (step 5120).

[0050] And the control block 430-2 of virtual-machine VM-2 of drawing 3 is specified, and OS is dispatched (step 5130). Then, transit (step 5140) of OS is started. While this OS runs, since the page fixed VM discernment bit 432 in virtual-machine VM-control block 430-2 of 2 is 1, conversion to the host absolute address from the guest absolute address is performed by adding the value of origin address field 433 to the guest absolute address. Furthermore, the upper limit check of the address is performed the value of the terminal point address field 434, and by comparing the host absolute address.

[0051] After transit initiation, since the immediate execution bit 431 of the control block 430-2 of VM-2 is 1 when OS on virtual-machine VM-2 publishes the starting initiation instruction of asynchronous process equipment 200 (step 5150), immediate execution of this instruction is carried out by the instruction processing unit 100, without being deduced by the virtual-machine monitor 410 (step 5160).

[0052] According to this example as mentioned above, with a central processing unit, a movable virtual machine system can be offered for a guest main storage area on main storage, without suspending access to the guest main storage area by the asynchronous process equipments 200 which process to asynchronous, such as a channel. Moreover, dynamically, the page immobilization VM can be changed into page possible VM, or page possible VM can be changed into the page immobilization VM.

[0053] The allotment modification approach of a computer resource is explained using allotment modification approach drawing 7 of II. computer resource, drawing 8, and drawing 9. Drawing 7 is the flow chart of the allotment modification approach of a calculating-machine resource, drawing 8 is the explanatory view of the modification approach of the size of a guest main storage area, and drawing 9 is the explanatory view of the configuration table 9000. The configuration table 9000 shows the computer resource assigned to each virtual machine an event generating front and after event generating. For example, the configuration table 9000 of drawing 9 is set to the initial state (condition before event generating) shown in drawing 8 (a). Virtual-machine VM-1 which has the guest main storage area 420-1 which consists of 32 megabytes, Virtual-machine VM-2 which have the guest main storage area 420-2 which consists of 128 megabytes, Three sets of the virtual machines of virtual-machine VM-3 which have the guest main storage area 420-1 which consists of 32 megabytes mean sharing and running a central processing unit 100-1 and 100-2.

[0054] Then, if a specific event occurs, the virtual-machine monitor 410 will demand OS on virtual-machine VM-2 to reduce to 32 megabytes specified as the configuration table for the guest main storage area 420-2 (step 7000). Here, a failure occurs to the event of having become specific time amount, and the 1st computer 500 shown in drawing 1, and a specific event shows the event that OS on virtual-machine VM-1 of the 2nd computer 600 succeeded processing of this 1st computer 500 with the system transfer device 700. Consequently, as shown in drawing 8 (b), OS on virtual-machine VM-2 makes the field of the high order address in guest main storage area 420-2 off-line (condition of not connecting) (step 7010), and separates it from virtual-machine VM-2.

[0055] Next, the virtual-machine monitor 410 moves like the case of the 1st example in the guest main storage area 420-2 of virtual-machine VM-2 (step 7020). By this, as shown in (c) of drawing 8, the field of the high order address contiguous to the guest main storage area 420-1 of virtual-machine VM-1 can be made into a free area.

[0056] Moreover, the virtual-machine monitor 410 demands OS on virtual-machine VM-1 to expand to 128 megabytes specified as the configuration table for the guest main storage area 420-1 (step 7030). Consequently, as shown in (d) of drawing 8, OS on virtual-machine VM-1 makes the field of the high order address contiguous to the guest main storage area 420-1 online (connection condition) (step 7040), and it connects with virtual-machine VM-1.

[0057] Furthermore, the virtual-machine monitor 410 makes virtual-machine VM-1 monopolize one set

of a central processing unit 100-1, makes a central processing unit 100-2 share with virtual-machine VM-2 between virtual-machine VM-3, and makes it run OS on a virtual machine, as shown in the configuration table 9000 (step 7050).

[0058] According to this example as mentioned above, the virtual machine system which can be changed dynamically can be offered for the size of a guest main storage area. Moreover, the virtual machine system which can be changed dynamically can be offered for allotment of the central processing unit to a virtual machine.

[0059]

[Effect of the Invention] According to this invention, the virtual machine system which can be changed dynamically can be offered for the size of a guest main storage area by making into a free area the field of an address high order which moves on main storage in the guest main storage area of each virtual machine, and adjoins the guest main storage area of a specific virtual machine.

[0060] Moreover, a movable virtual machine system can be offered for a guest main storage area on main storage, without suspending access to the guest main storage area according to asynchronous process equipment by moving one by one in a field from the field of the outside for [of the guest main storage area by asynchronous process equipments, such as a channel,] access.

[0061] Moreover, the virtual machine system which can be changed into page possible VM can be offered for the page immobilization VM by defining the response relation between the guest main storage area of the page immobilization VM, and the continuation field on the main storage with which this resides permanently as a host virtual space.

[0062] Moreover, the virtual machine system which can be changed into the page immobilization VM can be offered for page possible VM by matching the continuation field of a main storage area as a mapping place of a host virtual space.

[0063] Furthermore, the virtual machine system which can be changed dynamically can be offered for allotment of the central processing unit to a virtual machine by assigning the computer resource indicated by each virtual machine ignited by generating of an event at the configuration table

[Claim(s)]

[Claim 1] In the virtual machine system two or more operating systems (OS) in the calculating-machine top containing the central processing unit concerned and the asynchronous process equipment which processes to asynchronous run under control of a virtual-machine monitor according to the directions from a central processing unit, main storage, and said central processing unit While assigning the 1st fixed area of said main storage to the 1st OS among said two or more OS's The step which assigns the 2nd fixed area which adjoins the address high order side of said 1st fixed area to the 2nd OS, By moving the step which reduces the size of the 2nd fixed area assigned to said 2nd OS, and the 2nd fixed area which reduced this size to an address high order side on said main storage, when a specific event occurs The approach of having the step which generates a free area to the address high order side of said 1st fixed area, and the step to which the size of the main storage area of said 1st OS is expanded to assign the computer resource in a virtual machine system.

[Claim 2] The dynamic allocation approach of the computer resource in a virtual machine system of setting to claim 1 and having the step which defines the size of the main storage area of each OS generating before of said specific event, and after generating, and the step which changes the size of the main storage area of two or more of said OS's according to said definition when said specific event occurs.

[Claim 3] The dynamic allocation approach of the computer resource in a virtual machine system of having two or more said central processing units, and having the step which defines the central processing unit assigned to each OS the generating front of a specific event, and after generating, and the step which assigns a central processing unit to said two or more OS's according to said definition when said specific event occurs in claim 1.

[Claim 4] It is the dynamic allocation approach of the computer resource in a virtual machine system characterized by being the event that said specific event became specific time amount in claims 1, 2, or 3.

[Claim 5] It is the dynamic allocation approach of the computer resource in a virtual machine system which a failure occurs to the 2nd computer by which said specific event differs from said computer in claims 1, 2, or 3, and is characterized by being the event that said 1st OS succeeded processing of this 2nd computer.

[Claim 6] The step which generates said free area of claim 1 The step which considers as the virtual space where said virtual-machine monitor generated the main storage area of said 2nd OS, and maps this virtual space in said 2nd reduced fixed area, The dynamic allocation approach of the computer resource in a virtual machine system characterized by having the step which changes the mapping place of said virtual space into the field of the migration place of said main storage, and the step which assigns the field of this migration place as a fixed area of said 2nd OS.

[Claim 7] In claim 6 during migration of the main storage area of said 2nd OS The step to which said virtual-machine monitor carries out simulation of the beginning-of-using instruction of said asynchronous process equipment which said 2nd OS published, The step which analyzes which field said asynchronous process equipment accesses among the main storage areas of said 2nd OS at the time of this simulation, and the field used as this object for access The dynamic allocation approach of the computer resource in a virtual machine system characterized by having the step which moves after the access termination by said asynchronous process equipment.

[Claim 8] In the virtual machine system two or more operating systems (OS) in the calculating-machine top containing the central processing unit concerned and the asynchronous process equipment which processes to asynchronous run under control of a virtual-machine monitor according to the directions from a central processing unit, main storage, and said central processing unit When moving in the fixed area of the consecutive addresses of said main storage assigned as a main storage area of one of said OS's, The step which defines the response relation between the main storage area of the OS concerned, and said fixed area as the virtual space which said virtual-machine monitor generated, The dynamic allocation approach of the computer resource in a virtual machine system of having the step which changes the mapping place of said virtual space into the fixed area of the consecutive addresses of the

migration place on said main storage, and the step which assigns the fixed area of this migration place as a main storage area of said OS.

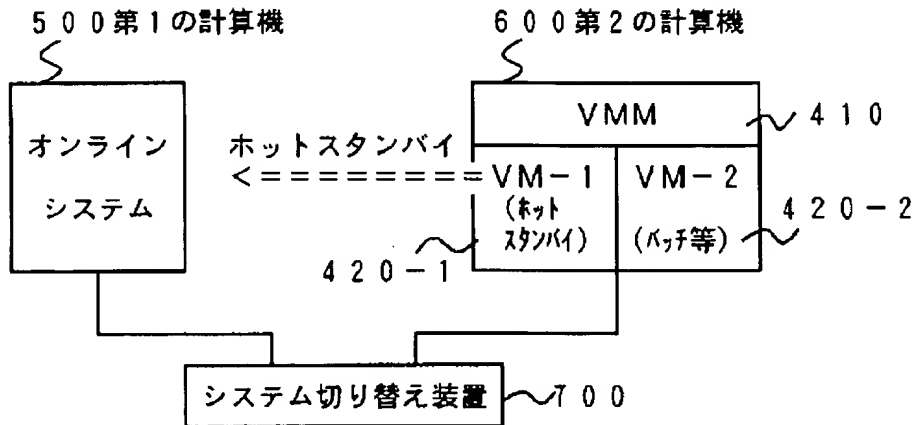
[Claim 9] The dynamic allocation approach of the computer resource in a virtual machine system characterized by having the step which makes size of the fixed area of the mapping place of said virtual space smaller than the size of the main storage area of said OS after said step to define in claim 8.

DRAWINGS

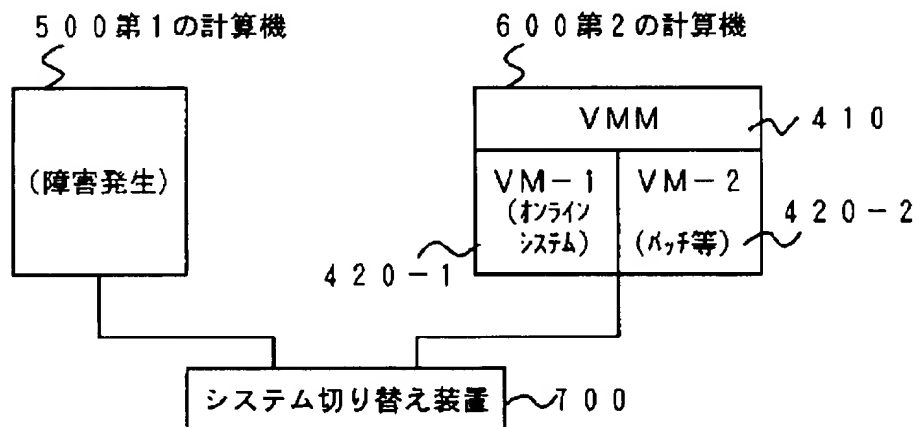
[Drawing 1]

ホットスタンバイシステムの説明図 (図1)

(a) システム切り替え前

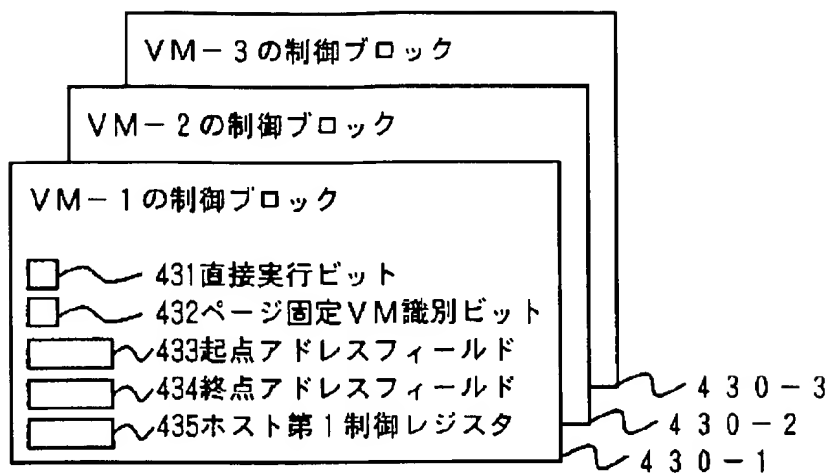


(b) システム切り替え後



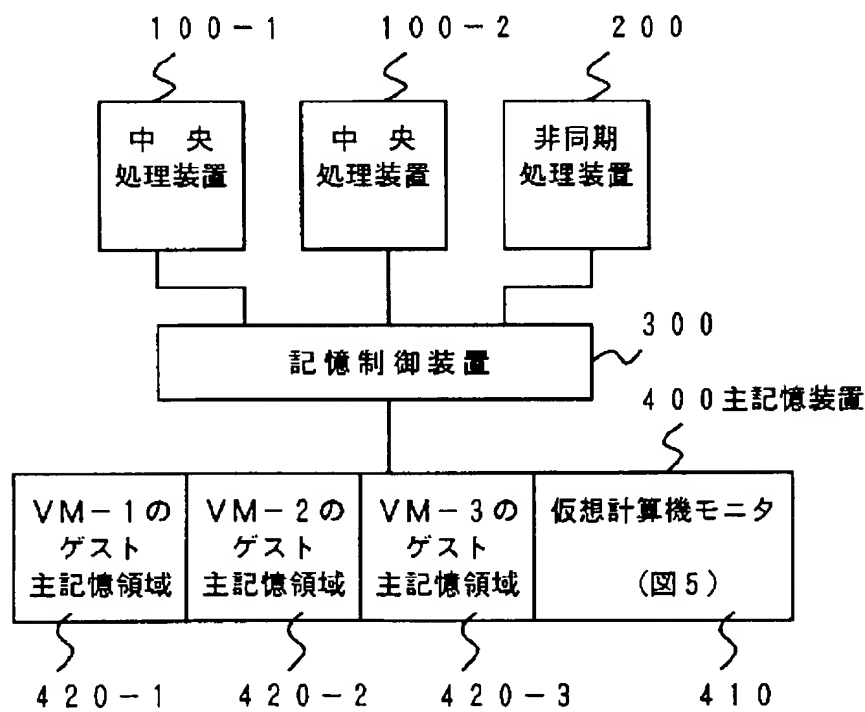
[Drawing 3]

仮想計算機（VM）の制御ブロックの説明図（図3）



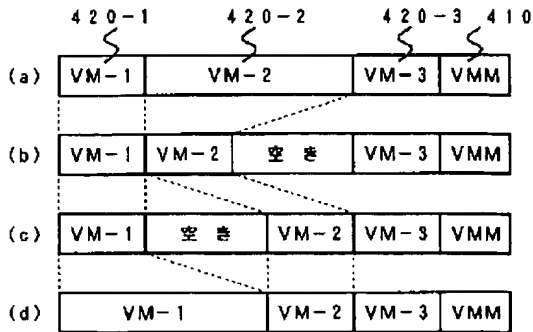
[Drawing 2]

仮想計算機システムの説明図（図2）



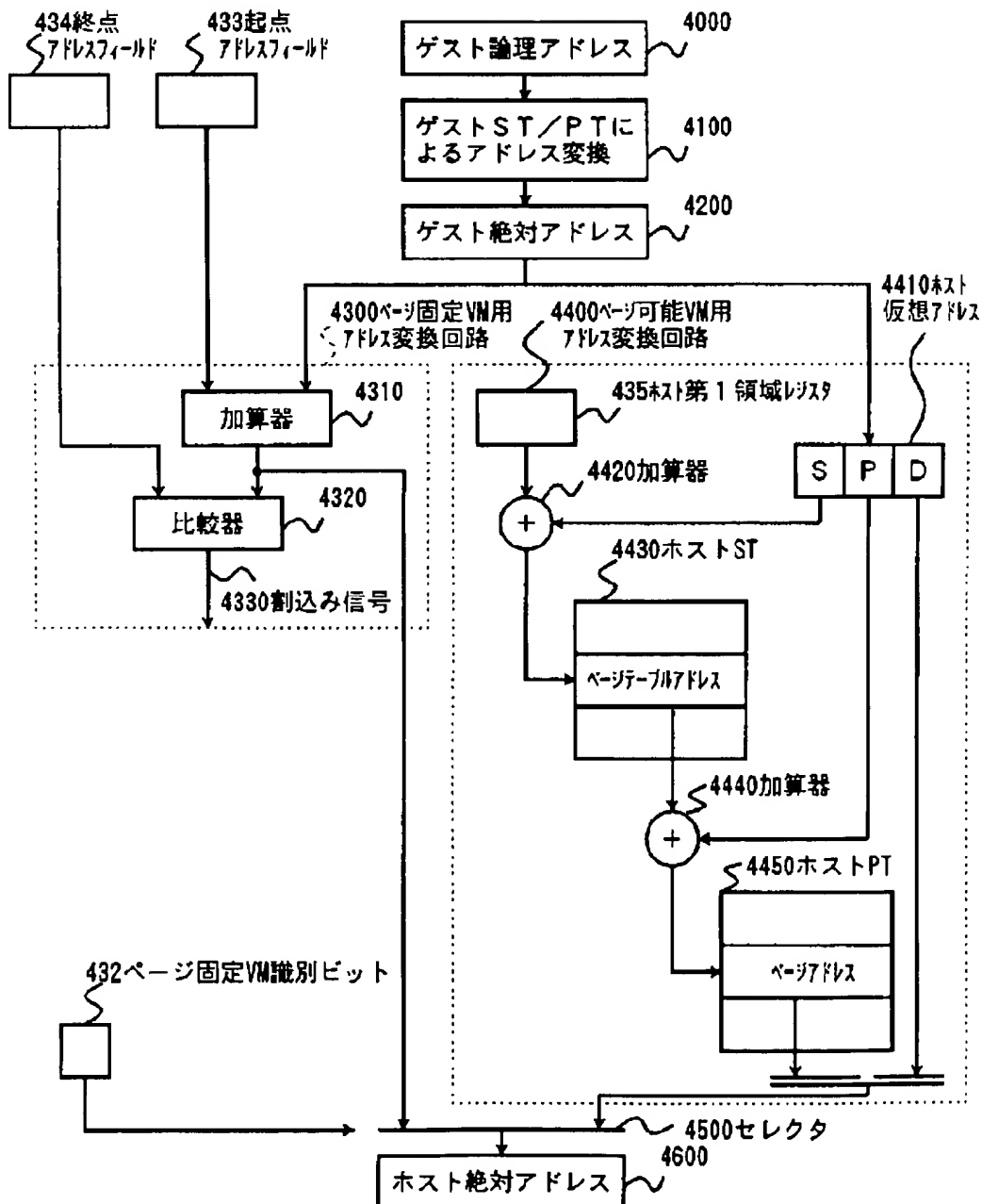
[Drawing 8]

ゲスト主記憶領域のサイズの変更方法の説明図 (図 8)



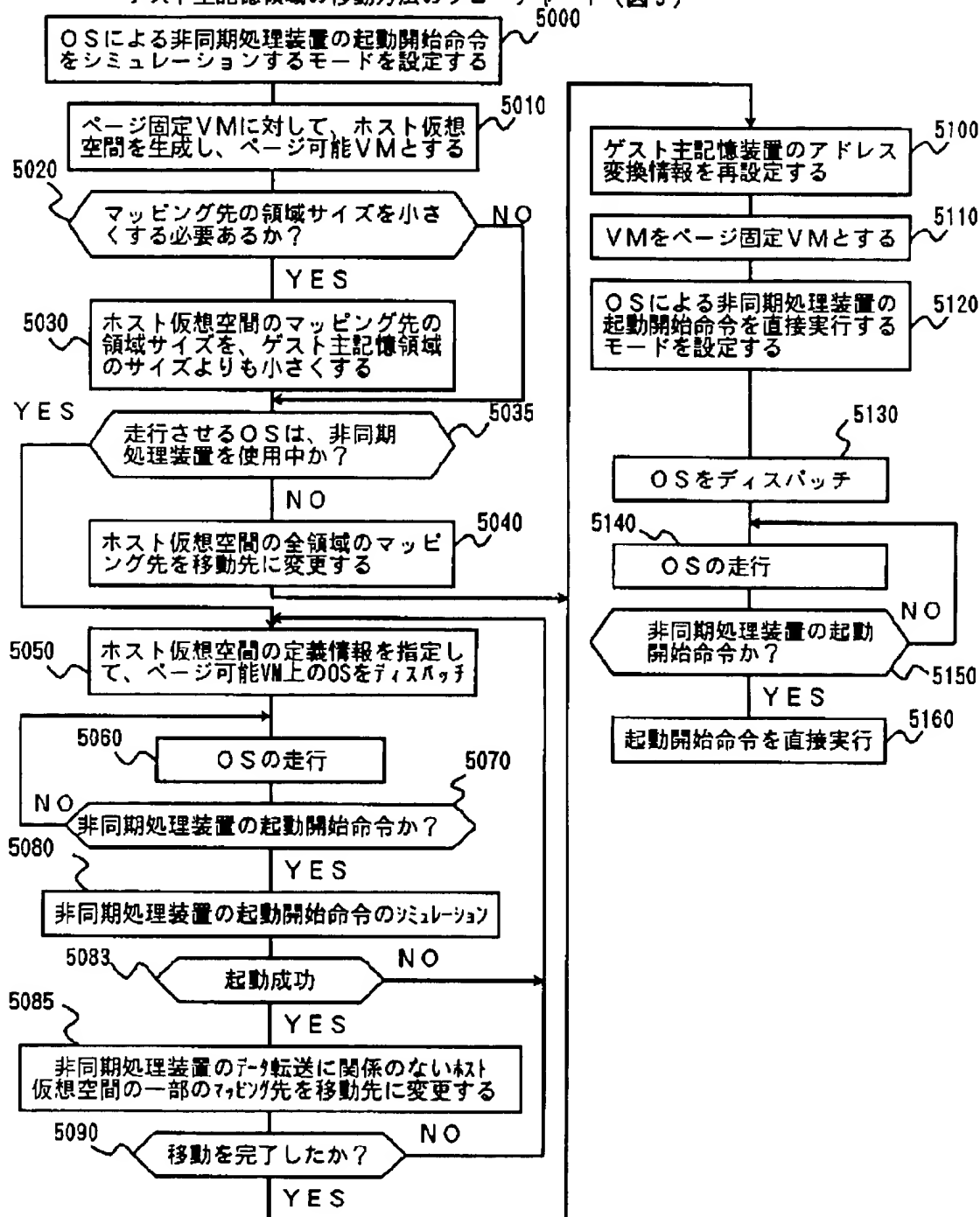
[Drawing 4]

仮想計算機システムにおけるアドレス変換の説明図 (図 4)



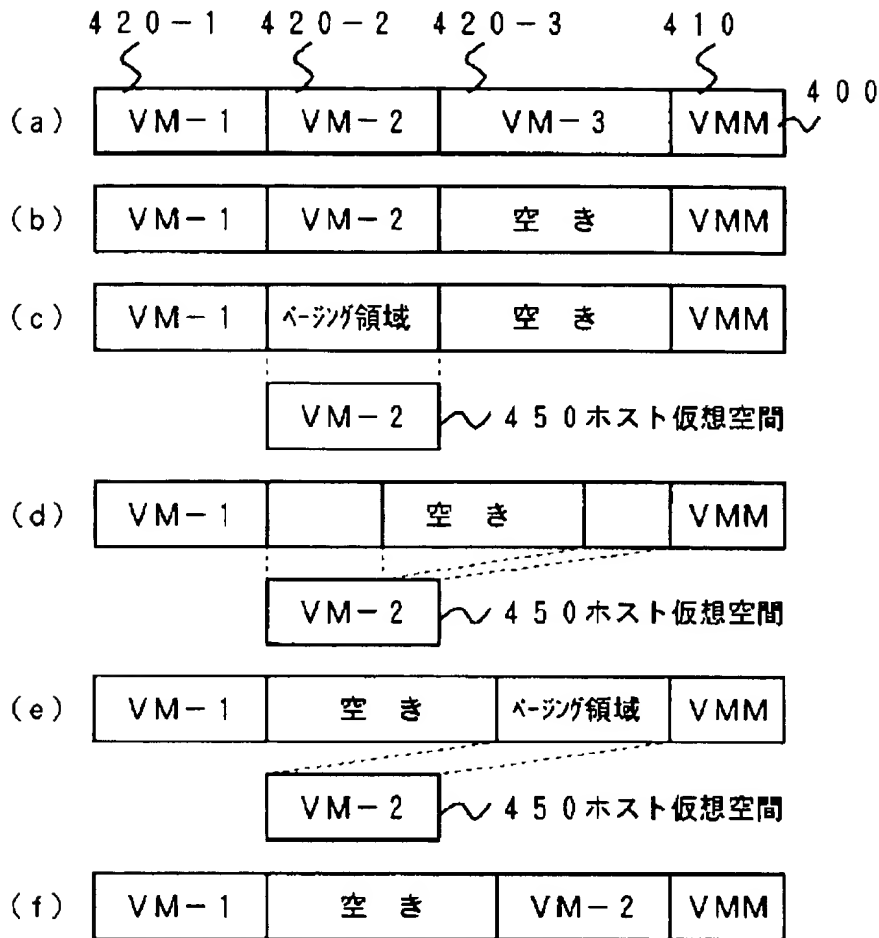
[Drawing 5]

ゲスト主記憶領域の移動方法のフローチャート (図5)



[Drawing 6]

ゲスト主記憶領域の移動方法の説明図 (図 6)



[Drawing 9]

構成テーブルの説明図 (図 9)

9 0 0 0 構成テーブル

		VM-1	VM-2	VM-3
イベント	発生前	ゲスト主記憶サイズ	3 2 MB	1 2 8 MB
	発生後	中央処理装置の割り当て	共 用	共 用
イベント	発生前	ゲスト主記憶サイズ	3 2 MB	3 2 MB
	発生後	中央処理装置の割り当て	専 有	共 用

[Drawing 7]

計算機資源の割り当て変更方法のフローチャート（図7）

